

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**International Publication No. WO 93/00437**

---

**Job No.: 1074-85208**

**Translated from French by the Ralph McElroy Translation Co.  
910 West Avenue, Austin, Texas, 78701**

**Ref.: 06832.0004-02000**

INTERNATIONAL PATENT OFFICE  
WORLD ORGANIZATION FOR INTELLECTUAL PROPERTY

International patent published on  
the basis of the Patent Cooperation Treaty (PCT)  
INTERNATIONAL PUBLICATION NO. WO 93/00437 A1

International Patent Classification <sup>5</sup> :	C 12 N      15/81
International Filing No.:	PCT/FR92/00559
International Filing Date:	June 19, 1992
International Publication Date:	January 7, 1993
Priority	
Date:	June 21, 1991
Country:	FR
No.:	91/07640

CLONING AND/OR EXPRESSION VECTORS, PREPARATION AND USE

Inventors; and	Xin-Jie Chen [CN/FR]
Inventors/Applicants (only for US):	15, rue Georges-Clémenceau Centre Universitaire Orsay F-91405 Orsay (FR)
	Reinhard Fleer [DE/FR] 47, avenue Beauséjour F-91440 Bures-sur-Yvette (FR)
	Hiroshi Fukuhara [JP/FR] 160, avenue du Général-Leclerc Résidence de Courcelles – Bât. 7 F-91190 Gif-sur-Yvette (FR)
Applicant (for all designated states except US):	Rhone-Poulenc Rorer S.A. [FR/FR] 20, avenue Raymond-Aron F-92160 Antony (FR)

Agent:

Philippe Becker  
Rhône-Poulenc Rorer S.A.  
Direction Brevets  
20, avenue Raymond-Aron  
F-92165 Antony Cédex (FR)

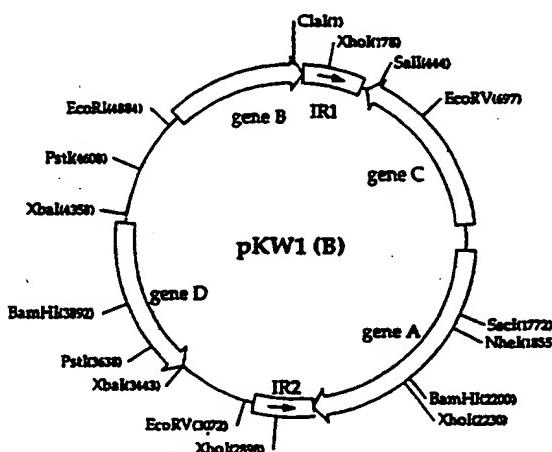
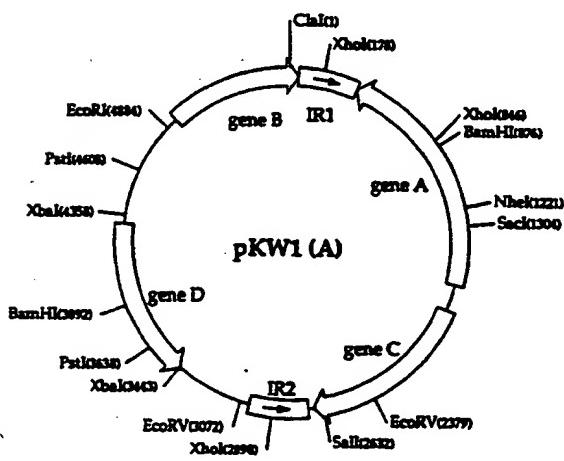
Designated States:

AU, CA, FI, HU, JP, NO, US,  
European Patent (AT, BE, CH, DE,  
DK, ES, FR, GB, GR, IT, LU, MC,  
NL, SE)

Published

With International Search Report.

Before expiration of the period permitted for amendments to the claims. Will be republished if amendments are submitted.



(57) Abstract

The invention discloses a new yeast plasmid, new cloning and/or expression vectors derived from said plasmid, and use thereof.

**FOR INFORMATION ONLY**

Codes for the identification of PCT contract states on the cover sheets of the documents that publish the international applications in accordance with the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MN	Mongolia
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GN	Guinea	NL	Netherlands
BJ	Benin	GR	Greece	NO	Norway
BR	Brazil	HU	Hungary	PL	Poland
CA	Canada	IE	Ireland	RO	Romania
CF	Central African Republic	IT	Italy	RU	Russian Federation
CG	Congo	JP	Japan	SD	Sudan
CH	Switzerland	KP	Democratic People's Republic of Korea	SE	Sweden
CI	Côte d'Ivoire	KR	Republic of Korea	SN	Senegal
CM	Cameroon	LI	Liechtenstein	SU	Soviet Union
CS	Czechoslovakia	LK	Sri Lanka	TD	Chad
DE	Germany	LU	Luxembourg	TG	Togo
DK	Denmark	MC	Monaco	US	United States of America

The present invention concerns a new yeast plasmid, as well as new cloning and/or expression vectors derived therefrom, their preparation and their use, notably for the production of recombinant proteins. The invention also concerns recombinant host cells containing such vectors.

More specifically, the present invention concerns all or part of a new plasmid isolated from the yeast *Kluyveromyces waltii* or a derivative thereof, and the cloning and/or expression vectors constructed from it.

During the past ten years, yeasts have become very promising host microorganisms for the production of heterologous proteins.

In particular, the identification by Beggs et al. (*Nature* 275 (1978) 104) of the plasmid  $2\mu$  and vectors derived therefrom was one of the keys of the current development of the genetic and molecular study of the yeast *Saccharomyces cerevisiae*. Since that time, the  $2\mu$  system has allowed the introduction and the expression of heterologous genes in yeast to obtain proteins of pharmaceutical or agro food interest. However, plasmid  $2\mu$  and its derivatives can replicate effectively only in yeasts belonging to the species *S. cerevisiae* and to some related species. Therefore, this system cannot be used for genetic manipulation in most yeast species. Notably, it cannot be used in yeasts whose physiological properties, which are different from those of *S. cerevisiae*, would allow better-performing new applications and/or industrial uses.

Different laboratories therefore have sought other plasmids in different yeast species. Thus, several circular plasmids have been found in the yeasts of the genus *Zygosaccharomyces*, and notably the plasmids pSR1 and pSR2 (Toh-é et al., *J. Bacteriol.* 151 (1982) 1380); pSB1, pSB2, pSB3 and pSB4 (Toh-é et al., *J. Gen. Microbiol.* 130 (1984) 2527); and pSM1 (Utatsu et al., *J. Bacteriol.* 169 (1987) 5537). A circular plasmid has also been found in *Kluyveromyces drosophilicola*: pKD1 (Falcone et al., *Plasmid* 15 (1986) 248).

All these plasmids present characteristics which they share with plasmid  $2\mu$ , notably inverted repeat sequences, and the existence of two possible isomeric forms due to a site-specific recombination site.

However, these plasmids always present the drawback of having a narrow host spectrum. Thus, because of their specificity with respect to yeast hosts, these plasmids can only be used in a limited number of strains.

The present invention is the result of the identification of a natural plasmid of the yeast *Kluyveromyces waltii*. This plasmid, called pKW1, is the first known natural plasmid of this yeast species. The plasmid has been purified from the strain *K. waltii* CBS 6430, and it has been mapped using restriction enzymes; the resulting map is presented in Figure 1.

An object of the invention thus consists of the plasmid pKW1 isolated from the strain *K. waltii* CBS 6430, or any fragment or derivative thereof.

In the meaning of the invention, the term derivative denotes plasmids which, in spite of some modifications, preserve the properties of the original plasmid. Notably, the modifications can take on the form of mutations or deletions concerning regions having a relatively large size. They can also be insertions or suppressions, for example, of cloning sites.

The term pKW1 fragment notably denotes the different genetic elements of this plasmid. It is more advantageous to notably mention, as genetic elements of pKW1, the structural genes or parts thereof, functional promoter sequences, inverted repeat sequences, or sequences (IR) that allow replication (replication origin) or confer stability to the plasmid (stability locus).

Indeed, the structural study of plasmid pKW1 has allowed the detection of analogies with the plasmid 2μ of *S. cerevisiae*. Thus, 4 structural genes have been identified (see Figure 2), as well as a replication origin. Moreover, the cloning of pKW1 in *E. coli* has allowed the isolation of 4 types of recombinant plasmids, corresponding to 2 isomeric forms of plasmid pKW1 (forms A and B) cloned in the vector pKan21 in the two possible orientations (vector pKan21 is described in Example 3.1). These two forms are presented in Figure 1. The existence of two isomeric forms indicates the presence of inverted repeat sequences. The study has thus shown that plasmid pKW1 contains a pair of inverted repeat sequences each having a length of 0.3 kb, and two unique sequences of 2.5 and 2.3 kb, whose orientation distinguishes the two isomeric forms A and B. The molecular size of plasmid pKW1 is thus approximately 5.5 kb. Different restriction sites have been detected, for example, the following sites: EcoRI, SphI, SalI, ClaI, NheI and BglII.

The complete nucleotide sequence of plasmid pKW1 has also been determined (Figure 3). The absence of homology between this sequence and that of known plasmids has been demonstrated, notably by hybridization experiments (see Example 2). The absence of hybridization under conditions of moderate stringency is characteristic of this difference in sequence.

In a preferred embodiment, the invention relates to a plasmid comprising all or part of the sequence presented in Figure 3 or a derivative thereof.

Moreover, the applicant has also shown that it is possible to use plasmid pKW1 or fragments thereof to construct particularly stable cloning and/or expression vectors.

Another object of the present invention thus consists of cloning and/or expression vectors characterized in that they comprise all or part of plasmid pKW1 of *K. waltii* CBS 6430 represented in Figure 1, or a derivative thereof.

A more specific object of the invention consists of a cloning and/or expression vector characterized in that it comprises at least one genetic element of plasmid pKW1.

Because of the host spectrum of plasmid pKW1, the vectors of the invention can be used in species other than the natural host *K. waltii*.

They can notably be used for the transformation of a large variety of species, notably yeast species.

Different types of vectors have been constructed from pKW1; they differ at the level of the size of the fragment originating from pKW1 and thus of the functional elements originating from pKW1.

An even more specific object of the invention consists of a cloning and/or expression vector characterized in that it comprises the replication origin of plasmid pKW1.

Other constructions can be prepared, containing fragments of varying sizes, to study the influence of the different elements of pKW1 on the stability of the vectors, their host specificity, and their effectiveness for the expression of the heterologous genes. In particular, expression vectors can be elaborated from the different genetic elements of plasmid pKW1 (replication origin, inverted repeat sequences, structural genes, promoter regions...), which one can introduce into known plasmids to improve their performances or to confer novel properties to them. In the same manner, vectors can be obtained by adding elements to plasmid pKW1, or by replacing certain genetic elements of pKW1 with elements originating from other plasmids. Thus, vectors can be obtained by substituting, for example, the replication origin of pKW1 with the replication origin of the plasmid 2μ of *S. cerevisiae* or of the plasmid pKD1 of *Kluyveromyces*, or by a chromosomal replicon (ARS) of yeast (example: KARS of *K. lactis*). In the same manner, vectors can be obtained by substituting the stability locus of pKW1 with that of plasmid 2μ of *S. cerevisiae* or plasmid pKD1 of *Kluyveromyces*. It can be particularly advantageous to prepare hybrid vectors which comprise plasmids pKD1 and pKW1.

Advantageously, the vectors of the invention comprise the entire plasmid pKW1 as represented in Figure 1.

It is preferred for the vectors according to the invention to comprise plasmid pKW1, linearized at a functionally neutral restriction site.

The term functionally neutral restriction site, in the meaning of the present invention, denotes a restriction site at whose level it is possible to interrupt the structure of the plasmid without altering its replication and stability properties.

In particular, the sites can be sites which are present on plasmid pKW1. As an example, one can notably mention the sites Clal(1); PstI(4608); or EcoRV(3072) as represented in Figure 1.

The sites can also have been artificially introduced into plasmid pKW1, or rendered unique. In this case, the sites are preferably introduced in intergenic regions of the plasmid, and notably in the region located between the genes B and D, or in the region located between gene D and IR2.

Advantageously, according to the present invention, plasmid pKW1 is linearized at the level of a unique restriction site.

A particularly interesting site in this regard is the unique *Cla*I site located in position 1 in Figure 1. Indeed, the applicant has shown that this site allowed the use of plasmid pKW1 to construct the cloning and/or expression vectors, by introducing at this level, for example, heterologous DNA fragments, while maintaining the stable replication of the vector obtained. This result is surprising to the extent that site *Cla*I is located in structural gene B.

The use of such neutral cloning sites thus allows the obtention of very stable vectors which are capable of staying in the transformed cells, even in the absence of any selective pressure.

Advantageously, the vectors of the invention contain, in addition, a heterologous DNA sequence comprising at least one structural gene, under the control of signals allowing its expression.

The signals allowing the expression of the structural gene(s) can consist of one or more elements chosen from promoters, terminators or secretion signals. It is understood that the signals are chosen as a function of the used host, the structural gene and the desired result. In particular, it can be preferable in some cases to use a promoter which can be regulated, allowing the decoupling between the growth phases of the hosts and the phase of expression of said structural gene(s). Similarly, the use of the peptide signal (secretion signal) can make it possible to increase the production rate of the desired protein and facilitate the purification step.

It is preferred for the promoters used to be derivatives of yeast genes. Promoters of particular interests are those derived from glycolytic genes of yeasts of the genus *Saccharomyces* or *Kluyveromyces*. Notably, one can cite the promoters of genes coding for phosphoglycerate kinase of *S. cerevisiae* (PGK), glyceraldehyde-3-phosphate dehydrogenase (GPD), enolases (ENO), alcohol dehydrogenases (ADH). One can also cite promoters derived from strongly expressed genes, such as the lactase gene (LAC4) or the acid phosphatase gene (PHO5).

Moreover, these genes can be modified by mutagenesis, for example, to add additional transcription control elements, such as, notably, UAS ("Upstream Activating Sequence") regions.

The structural gene which can be introduced into the vectors of the invention preferably codes for a polypeptide of pharmaceutical or agro food interest. For example, one can mention enzymes, such as, notably, superoxide dismutase, catalase, amylases, lipases, amidases, chymosine, etc.), blood derivatives (such as serum albumin, alpha- or beta-globin, factor VIII, factor IX, van Willebrand factor, fibronectin, alpha-1 antitrypsin, etc.), insulin and its variants, lymphokines (such as interleukines, interferons, colony stimulation factors (G-CSF, GM-CSF, M-CSF...), TNF, TRF, etc.), growth factors (such as growth hormone, erythropoietin, FGF,

EGF, PDGF, TGF, etc.), apolipoproteins, or antigenic polypeptides for the manufacture of vaccines (hepatitis, cytomegalovirus, Epstein-Barr, herpes, etc.).

In a special embodiment of the invention, the structural gene can be a gene resulting from the fusion of several DNA sequences. The gene can notably be a gene coding for a hybrid polypeptide containing an active part associated with the stabilizing part. As an example, one can mention the fusion between albumin or albumin fragments and a receptor part of a virus receptor (CD4, etc.).

In another embodiment, the heterologous DNA sequence can comprise several structural genes, notably genes involved, at the genetic or biochemical level, in the biosynthesis of a metabolite. The metabolite can be, in particular, an antibiotic, an amino acid or a vitamin.

In a special embodiment, the vectors of the invention contain, in addition:

- an *E. coli* replicon and/or
- at least one selection marker.

These elements allow the manipulation of the vectors of the invention in a much easier manner.

Another object of the invention concerns recombinant cells containing a vector as defined above.

The recombinant cells are preferably chosen from yeasts.

The applicant has shown that the vectors of the invention can indeed be used either in *K. waltii* (the natural host of pKW1) or in yeasts of different species or even different genera. Notably, they can be used in other *Kluyveromyces* species or in *Saccharomyces*. Moreover, when strain *K. waltii* CBS 6430 is used as the host cell, homologous recombinations between the vectors of the invention and the resident plasmid pKW1 can affect the stability of the vectors, and thus decrease the performances of the host/vector pair. To further improve the stability of the vectors of the invention in such a host/vector pair, the applicant has prepared a strain *K. waltii* pKW1<sup>+</sup> (KW18). This strain allows the optimization of the industrial use of the vectors of the invention (see Example 4).

Different techniques can be used to introduce the vectors of the invention into the host cells. In particular, transformation (Bianchi et al., *Curr. Genet.* 12 (1978) 185) and electroporation (Delorme, *Appl. Environ. Microbiol.* 155 (1989) 2242) yield good results. However, it is clear that the invention is not limited to a particular technique.

An object of the invention also consists of a method for the preparation of a polypeptide according to which a recombinant cell as defined above can be cultured and the produced polypeptide recovered. More specifically, the method of the invention allows the production of proteins of pharmaceutical or agro food interest, such as those indicated above. More

specifically, the method of the invention can be adapted to the production of human albumin and of its variants or precursors.

In the case where the structural genes are involved in the biosynthesis of a metabolite, the recombinant cells can also be used directly in a bioconversion process.

Other advantages of the invention will become apparent after a reading of the following examples which must be considered to be illustrative and nonlimiting.

#### Legend of the figures

Figure 1: Restriction map of plasmid pKW1. The inverted repeat sequences as well as the structural genes A-D are indicated. The positions indicated for the restriction sites correspond to the first nucleotide recognized by the enzyme.

Figure 2: Study of the open frames of plasmid pKW1. The genetic elements indicated are localized at the following positions with reference to the sequence presented in Figure 3: Gene A: nucleotides 1454-2755; gene B: nucleotides 4948-54; gene C: nucleotides 389-1309 on the complementary strand; gene D: nucleotides 3444-4313 on the complementary strand; IR1: nucleotides 53-368; IR2: nucleotides 2713-3028.

Figure 3: Nucleotide sequence of plasmid pKW1. The sequence represented corresponds to the B form of the plasmid. Position 1 corresponds to the first nucleotide of the sequence recognized by the enzyme Clal. Plasmid pKW1 was fragmented by restriction enzymes, and the fragments were cloned in pTZ18R (Pharmacia). The sequences of the cloned segments were determined using the 2-strand Sanger method.

Figure 4: Restriction maps of the shuttle plasmids YIP5 and pKan21. Ap: ampicillin resistance gene; Tc: tetracycline resistance gene; Km: kanamycin resistance gene (G418); LacZ: structural gene of  $\beta$ -galactosidase.

Figure 5: Strategy of construction of the vectors pBNA1, pNEA2, pBNB1/A3, pNEB1, pXXY2 and pXXK3. See also Table 1.

Figure 6: Restriction map of the vector pXXK3.

Figure 7: Strategy of construction of the vectors pKWC11, pKWS1 and pKWS14.

Figure 8: Study of the stability of vectors pKWC11 and pXXK3 in the strain *K. waltii* KW18. In each case, a transformed clone was cultured in nonselective YPG medium for the indicated number of generations, and then aliquots of the culture were plated on YPG agar Petri dishes with and without G418, to determine the total number of cells and the number of cells resistant to G418. The stability corresponds to the % of resistant cells.

Figure 9: Strategy of construction of the expression vector pXPHO5. Abbreviations: P = promoter, T = terminator, ss = secretion signal, CIP = Calf Intestinal Phosphatase, Km = kanamycin, E = EcoRI, H = HindIII, S = SalI, B = BamHI, Sm = SmaI.

Figure 10: Restriction maps of the vectors pXKN18 and pXPHO5. Legend: see Figure 9.

Figure 11: Immunological detection of IL1 $\beta$  produced by *K. waltii*. The molecular weight markers (KDa) are indicated on the left. Well 1: Reference IL-1 $\beta$  (100 ng); well 2: culture supernatant of the transformant pXKN18 (without IL-1 $\beta$  cassette); well 3: culture supernatant of the transformant pXPHO5 treated with endo-N-acetylglucosaminidase H; wells 4 and 5: supernatants of the cultures of the transformants pXPHO5 in LPi and HPi medium, respectively.

Figure 12: Strategy of construction of the plasmid pYG65.

Figure 13: Strategy of construction of the plasmid pYG70.

Figure 14: Strategy of construction of the plasmid pYG141. aph: gene coding for aminoglycoside 3'-phosphotransferase, which confers resistance to kanamycin; bla: gene coding for the  $\beta$ -lactamase which confers resistance to ampicillin.

Figure 15: Strategy of construction of the plasmid pYG142.

Tables 1 and 2: Composition of vectors derived from pKW1 according to the invention.

#### (\*) Cloning sites.

Tables 3, 4 and 5: Transformation by the vectors of the invention of *S. cerevisiae*, *K. waltii* and different strains of *Kluyveromyces*, respectively. The stability of the transformants is expressed by the percentages of Ura $+$  cells after 10 generations of growth in nonselective YPD medium. The isonuclear strains *K. waltii* pKW1 $^+$  and pKW1 $^-$  are CBS 6430 and KW18, respectively.

#### General cloning technique

The classical methods of molecular biology, such as centrifugation of plasmid DNA in a cesium chloride-ethidium bromide gradient, digestion by restriction enzymes, gel electrophoresis, electroelution of DNA fragments from agar gels, transformation in *E. coli*, etc., are described in the literature (Maniatis et al., "Molecular Cloning: a Laboratory Manual," Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1986; Ausubel et al. (eds.), "Current Protocols in Molecular Biology," John Wiley & Sons, New York 1987).

In vitro mutagenesis directed by oligodeoxynucleotides is carried out according to the method developed by Taylor et al. (Nucleic Acids Res. 13 (1985) 8749-8764) using the kit distributed by Amersham. Sequencing of nucleotides is carried out according to the dideoxy technique described by Sanger et al. (Proc. Natl. Acad. Sci. USA 74 (1977) 5463-5467). Enzymatic amplification of specific DNA fragments is carried out by the PCR ("Polymerase-catalyzed Chain Reaction") reaction under the conditions described by Mullis and Falloona (Meth. Enzym., 155 (1987) 335-350) and Saiki et al. (Science 230 (1985) 1350-1354), using a "DNA thermal cycler" (Perkin-Elmer Cetus) following the recommendations of the manufacturer.

## Examples

### 1) Isolation and purification of pKW1

Strain CBS 6430 is cultured in 2 L of YPG medium (yeast extract 1%, Bactopeptone 1%, glucose 2%) with stirring at 26°C for approximately 18 h. The cells in the early stationary phase are harvested by centrifugation. 13-15 g of cell mass per liter are usually obtained. The cells are washed with 150 mL of 1M sorbitol containing 30 mg of zymolyase 20T (Kirin Breweries Co., Tokyo). After incubation at 30°C for 1 h, one adds to the suspension of protoplasts 5 mL of 10% sodium dodecyl sulfate and 5 mL of 0.5M EDTA, pH 7.0. The mixture is immediately vigorously stirred and incubated at 50°C for 1-2 h. Potassium acetate is added to the lysate to obtain a final concentration of 1M and the mixture is maintained in ice for 2 h. The precipitates formed are eliminated by centrifugation (Sorvall SS34, 15,000 rpm, 30 min). The supernatant, to which 2 volumes of 95% ethanol are added, is cooled in ice to precipitate the nucleic acids. The precipitates are collected by centrifugation, washed with ethanol at 70%, dried under a vacuum, and finely dissolved in 40 mL of 5 x TE (1x TE is 10mM tris-HCl, 1mM EDTA, pH 8). After the addition of 40 g of CsCl in 5 mL of ethidium bromide (stock solution of 10 mg/mL), the mixture is centrifuged at 60,000 rpm for 6 h (Beckman, rotor 60 Ti). The fluorescent band of the plasmid DNA is located below the major band of the chromosomal DNA. The DNA of the plasmid is collected and subjected to the second centrifugation cycle in CsCl/ethidium bromide. The collected plasmid DNA is mixed with one volume of isopropanol which has first been equilibrated with 4M CsCl to eliminate the ethidium bromide. After several extractions with isopropanol, the DNA solution is dialyzed against 1 x TE. The volume of the solution can be reduced by dialysis against polyethylene glycol 6000 in flakes.

### 2) Sequencing and studies of homologies

Plasmid pKW1 was sequenced using the method described by Sanger et al. (Proc. Natl. Acad. Sci. USA 74 (1977) 5463-5467). The complete sequence is indicated in Figure 3.

The absence of a homology between this sequence and that of known plasmids was determined by molecular hybridization experiments: pKW1 is labeled with  $^{32}\text{P}$  and hybridized with the following plasmids, which were first immobilized on a nitrocellulose filter:

- plasmid 2 $\mu$  from *S. cerevisiae*,
- pSR1, pSB3 and pSB4 from *Z. rouxii*,
- pSB1 and pSB2 from *Z. bailii*, and
- pKD1 from *K. drosophilaeum*.

The hybridization was carried out under moderately stringent conditions (0.6M Na $^+$ , 65°C, 18 h); none of the plasmids produced a positive signal of hybridization.

The other circular plasmids (pSB1, pSB4) whose sequences have not yet been described are also different from pKW1, in the molecular size, the length of the inverted repeats, and in their host species.

### 3) Constructions of cloning vectors derived from pKW1

Two types of recombinant molecules were constructed from pKW1.

3.1 In the first type, different fragments of pKW1 (corresponding, for example, to the genetic elements of pKW1) were introduced into shuttle vectors, and notably into the vectors YIp5 (Struhl et al., Proc. Nat. Acad. Sci USA 76 (1979) 1035) and pKan21, represented in Figure 4.

Plasmid pKan21 was constructed by insertion of the gene *aph* (Genblock, Pharmacia) which confers resistance to kanamycin (G418) in the form of a 1.25 kbp AccI fragment, in the NarI site of plasmid pUC19 (Viera and Messing, Gene 19 (1982) 259). Thus, pKan21 contains, in addition to *aph*, the gene *bla* which confers resistance to ampicillin and the replication origin of ColE1 which allows replication in *E. coli*. The yeasts transformed by the vectors derived from pKan21 can be detected by their growth on a medium containing 200 µg/mL of geneticine (G418).

Plasmid YIp5 is a derivative of plasmid pBR322 in which the gene URA3 from *S. cerevisiae* was inserted as a selection marker. The yeast, transformed by the derivatives of YIp5, is detected by its growth on a uracil-free medium. In the last case, the host yeast is an auxotroph which is deficient in orotidine-monophosphate carboxylase of the uracil synthesis pathway.

The vectors of this first type are pBNA1, pNEA2, pBNB1/A3, pNEB1, pXXY2 and pXXK3 (Table 1 and Figure 4).

#### - Vector pBNA1

pKW1 is digested with BglII and NheI (see Figure 5) and the DNA fragments are separated by electrophoresis. The 2.4-kbp fragment is recovered and inserted by ligation between the unique BamHI and NheI sites in the tetracycline resistance gene of plasmid YIp5 (Yanish-Peron et al., Gene 33, 1985, 103-119), where the latter was first digested by BamHI and NheI and repurified. The BglII and BamHI ends are compatible for a specific ligation. *E. coli*, transformed by the ligation mixture, is selected on LB agar medium which contains ampicillin. The insertion is verified by replication of the transformants on LB agar medium containing tetracycline, as the latter are sensitive to this antibiotic. The structure of the plasmid is verified by extraction of the DNA from individual transformants and its analysis using restriction enzymes. For example, plasmid pBNA1, digested by PstI, produces 3 fragments of 3.28, 3.15 and 1.36 kbp: double digestion with PstI/NheI yields 4 fragments of 3.15, 2.3, 1.36 and 1.0 kbp.

- Vector pNEA2

The 2.0 kbp NheI-EcoRI fragment of pKW1 (see Figure 5) is isolated and inserted between the two unique NheI and EcoRI sites in the tetracycline resistance gene of YIp5, where the latter was first digested by these two enzymes. The ligation product is introduced into *E. coli* and the transformants which are resistant to ampicillin and sensitive to tetracycline are isolated. The plasmid is isolated, as in the case of pBNA1, from one of its transformants. The structure of the recombinant plasmid obtained is verified by restriction. For example, digestion by PstI produces 2 fragments of 4.35 and 3.15 kbp; double digestion with PstI/XbaI produces 4 fragments of 3.15, 1.8, 1.75 and 0.7 kbp.

- Vector pXXY2

The 545 bp XbaI-XbaI fragment of pKW1 (see Figure 5) is isolated and inserted between the unique SalI-NheI sites in the tetracycline resistance gene of plasmid YIp5. The XbaI and SalI ends, on the one hand, and XbaI and NheI ends, on the other hand, are compatible for a specific ligation. The ligation product is introduced into *E. coli*, and the recombinant plasmid is isolated as in the case of pBNA1. The structure of the plasmid obtained is verified by restriction. For example, digestion with EcoRI + NruI produces two fragments of 4.5 and 0.96 kbp.

- Vector pXXK3

The 545 bp XbaI-XbaI fragment of pKW1 (see Figure 5) is isolated and inserted between the unique SalI-XbaI sites (polylinker at the N terminal of LacZ) of plasmid pKan21. The ligation mixture is introduced into *E. coli*, and the transformants, which are plated on LB medium containing X-gal and IPTG, are isolated as white colonies among blue colonies. When replicated on LB medium containing kanamycin, they grow well. These colonies are individually analyzed for their plasmid content. The plasmid obtained from one of the transformants presents the structure shown in Figure 6, which was verified by restriction. For example, digestion with BamHI + PstI produces four fragments: 2.5, 1.3, 0.55 and 0.19 kbp.

- Vector pBNB1/A3

The 1.9 kbp BglII-NheI fragment of pKW1 (see Figure 5) is isolated and inserted between the BamHI-NheI sites of YIp5. The recombinant plasmid is isolated as in the case of pBNA1. The structure of the plasmid is verified by restriction. For example, digestion with PstI produces 3 fragments of 3.15, 2.8 and 1.36 kbp; double digestion with PstI/NheI produces 4 fragments of 3.15, 1.8, 1.36 and 1.0 kbp.

- Vector pNEB1

The 2.5 kbp NheI-EcoRI fragment of pKW1 (see Figure 5) is isolated and inserted between the NheI-EcoRI sites of YIp5. The recombinant plasmid is isolated as in the case of pBNA1. The structure of the plasmid is verified by restriction. For example, digestion with PstI produces 2 fragments of 4.65 and 3.15 kbp; double digestion with PstI/XhoI produces 3 fragments of 3.15, 2.9 and 1.75 kbp.

3.2 The second type of recombinant molecules contains the totality of the sequence of pKW1. To obtain these vectors, pKW1 is linearized by a unique cut at a restriction site, allowing the introduction of heterologous DNA segments. Such segments can contain structural genes, included, for example, in expression cassettes and/or whole shuttle vectors, such as, notably, pKan21 or YIp5 (Figure 4).

Examples of this type of vectors are pKWC11, pKWS1 and pKWS14 (Table 2 and Figure 7).

- Vector pKWS14

The DNA of pKW1 is digested by the restriction enzyme SalI. Plasmid pKan21, described above, is also digested by SalI (the unique SalI site is localized in the cloning multisite present in the LacZ gene). The two plasmids are ligated by DNA ligase. The ligation mixture is used for the transformation of *E. coli* JM83 as in the preceding case. The suspension of the cells transformed is plated on LB agar containing X-gal and IPTG. The white colonies, among the blue ones, are harvested individually. They are resistant to ampicillin and kanamycin. Their plasmid content is analyzed on DNA minipreparations as above. Plasmid pKWS14, isolated from one of the transformants, contains form A of plasmid pKW1 and it has the structure shown in Figure 7. It is verified by restriction. For example, digestion with BamHI produces 3 fragments of 5.6, 2.6 and 1.2 kbp.

- Vector pKWC11

Plasmid pKW1 is digested with ClaI. Plasmid pKan21 is digested with AccI. They are repurified by treatment with phenol and precipitation with ethanol. The DNA of the two plasmids are mixed in approximately equal quantities and subjected to ligation with DNA ligase for one night. The ligation product is amplified in *E. coli*. The colonies of the transformants are white on LB medium containing Xgal and IPTG. They are resistant to kanamycin and ampicillin in the medium containing one or the other of these antibiotics. Plasmid pKWC11, isolated from one of the transformants, contains form A of plasmid pKW1 and has the structure shown in Figure 7. It is verified by restriction. For example, digestion of the plasmid with BamHI produces 3 fragments of 5.6, 2.9 and 0.9 kbp.

### - Vector pKWS1

Plasmid pKW1 and plasmid YIp5 are digested with Sall. The mixture is repurified and subjected to ligation. The ligation product is introduced into *E. coli*. Transformants which are resistant to ampicillin and sensitive to tetracycline are obtained. Plasmid pKWS1, isolated from one of them, contains form A of plasmid pKW1 and it has the structure shown in Figure 7. For example, digestion of the plasmid with EcoRI yields two fragments of 8.3 and 2.7 kbp.

### 4) Construction of a strain CBS 6430 pKW1

*K. waltii* CBS 6430 was first transformed by the recombinant plasmid pKWS14 (Table 2 and Figure 7). The transformation method used is essentially that described by Chen and Fukuhara (Gene 69, 181 (1988)) using protoplasts. The transformants obtained are maintained for 75 generations on YPD agar medium 1 mg/mL G418 (YPD medium: yeast extract 10 g/L; peptone 20 g/L; glucose 20 g/L). They are then transferred to a liquid YPD medium without antibiotic, and maintained for 10 generations.

Under these conditions (without selective pressure), plasmid pKWS14 is progressively lost. The colonies which are sensitive to G418 and which appear are collected and individually tested for the presence of the plasmids. The test consists of the extraction of cellular DNA, followed by electrophoresis of these DNAs on agar gel. The presence of plasmids is determined by staining with ethidium bromide. Among the colonies which have become sensitive to G418, 25% were shown to lack any plasmid. One of these colonies was kept as a strain of *K. waltii* which lacked plasmid pKW1, and it was called KW18.

### 5) Transformation of different yeasts

#### 5.1 Transformation of *S. cerevisiae*

Among the different vectors described in Tables 1 and 2 and in Figures 4 and 5, some of those containing the marker URA3 were used to transform an auxotrophic strain ura3 of *S. cerevisiae* (strain S150-2B: Mat a, ura3, leu2, trp1, his3, 2 $\mu$ ).

The transformation method is essentially the one described by Sherman et al. (Yeast Genetics, Cold Spring Harbor, NY, 1986).

The results obtained are indicated in Table 3. They show that the vectors of the invention are capable of transforming the yeasts of the genus *Saccharomyces*.

#### 5.2 Transformation of *K. waltii*

The transformation of *K. waltii* was carried out with vectors which carry the marker for kanamycin resistance.

The transformation method used is essentially the one described by Chen and Fukuhara (Gene 69 (1988) 181) using protoplasts. It is clear that any other technique which allows the introduction of a DNA fragment in a microorganism can be used.

The results obtained are presented in Table 4.

They show that the vectors of the invention are capable of transforming the yeast *K. waltii* with a high frequency.

Moreover, the stability study described in Figure 8 shows that vectors can be obtained from pKW1 presenting a stability of 100% after 50 generations of growth in a nonselective medium. This is perfectly illustrated by vector pKWC11. This study also shows that it is preferable, to obtain a relatively high stability, to use the vectors which comprise only the replication origin of plasmid pKW1 in host cells having a resident pKW1 plasmid.

### 5.3 Transformation of other yeasts

Vector pKWC11, which is highly stable and autonomous in *K. waltii*, was used to test the transformability of different yeast species, notably those belonging to the genus *Kluyveromyces*.

The results are presented in Table 5.

The presence of vector pKWC11 in the transformants was verified by electrophoresis.

All of these results show that the range of host species for the vectors of the invention can be very large, beyond the genus *Kluyveromyces*.

## 6) Use of the vectors of the invention for the production of heterologous proteins

### 6.1 Interleukin-1 $\beta$ :

6.1.1 Construction of an expression and secretion vector of IL-1 $\beta$  derived from pKW1 (Figures 9 and 10).

- Vector pXXK3 (Table 1, Figure 6) is linearized by EcoRI, and the ends are filled with the Klenow fragment of DNA polymerase I of *E. coli*. A synthetic "linker" (5'-GC GGCCGC-3') forming a restriction site which is recognized by the enzyme NotI is added by means of T4 ligase, and the vector so obtained (pXKN18) is purified after its amplification in *E. coli* (Figure 10).

- An expression cassette of IL-1 $\beta$  is prepared, consisting (a) of the regulated promoter PHO5 originating from *S. cerevisiae* (Bajwa et al., Nucl. Acid. Res. 12 (1984) 7721-7739), (b) the gene of human IL-1 $\beta$  (Jung et al., Ann. Inst. Pasteur/Microbiol. 139 (1988) 129-146) preceded (c) by a synthetic sequence corresponding to the signal sequence of the killer toxin of pGKL1 of *K. lactis* (pre region of the gene of the alpha subunit) (Stark and Boyd, EMBO J. 5, (1986) 1995-2002), and (d) by the terminator PHO5. The expression cassette was isolated from the vector pSPHO5-IL14 whose construction is described in the patent EP 361991. The cassette

was prepared as follows: At the 5' end of the gene coding for the mature part of IL-1 $\beta$ , the following synthetic sequence is inserted, in the form of an EcoRI fragment:

Met Asn Ile Phe Tyr Ile Phe Leu Phe Leu Ser Phe Val Gln Gly Lys Arg  
 5'-AATTATGAATATAATTACATATTTTGTGTCATTGTTCAAGGTAAAAG-3'  
 3'-TACTTATATAAAATGTATAAAAACAAAAACGACAGTAAGCAAGTTCCATTTCCTTAA-5'

The last codons added (Lys and Arg) form a potential restriction site which is recognized by the endopeptidase Kex1 of *K. lactis* (Tanguy-Rougeau et al.; FEBS Lett. 234 (1988) 464). This sequence was fused to the IL-1 $\beta$  gene by the EcoRI site, forming the following junction:

Gly Lys Arg Ile His Met Ala  
 5'....GGT AAA AGA ATT CAT ATG GCA ....3'

Alanine (GCA) corresponds to the first amino acid of mature IL-1 $\beta$ . Arg-Ile-His-Met corresponds to an EcoRI-NdeI "linker" introduced to facilitate the cloning (see EP 361 991).

The entire cassette is put in the form of a NotI fragment by the addition of a corresponding linker (5'-GC<sub>n</sub>GC<sub>n</sub>GC-3').

- The secretion cassette of IL-1 $\beta$  is inserted, at the NotI site, in pXKN18. The resulting vector is called pXPHO5 (Figure 10).

**6.1.2** The strain *K. waltii* CBS 6430 is transformed with vector pXPHO5, under the conditions described in Example 5.2.

#### 6.1.3 Expression of IL-1 $\beta$ :

The transformed cells are cultured at 28°C, in the absence of G418, in the liquid medium LPi (low mineral phosphate content) and the medium HPi (high phosphate content), prepared according to Chen and Fukuhara (Gene 69 (1988) 181-192), for 4 days. 50 mL of culture are centrifuged and the supernatants are filtered through a Millipore membrane (0.22  $\mu$ m). The proteins are precipitated by the addition of ethanol at the final concentration of 60%. The precipitates are dissolved in 2 mL of Laemmli buffer (Nature 227 (1970) pp. 680-685) and 20- $\mu$ L samples are used for SDS-PAGE analysis according to Laemmli (document cited above). After the electrophoresis, the proteins are transferred to a nitrocellulose sheet, and treated with a polyclonal anti-human IL-1 $\beta$  rabbit antiserum. The blot is then treated with a second biotinylated

anti-rabbit polyclonal antibody (Vectastain ABC ImmunoPeroxidase Kit, Vector Laboratories). The antigen-antibody complex is developed according to the protocol of the supplier.

Figure 11 shows that a protein having an apparent molecular weight of 21 kDa is secreted by the yeast transformed by pXPHO5. The protein is specifically recognized by the anti-IL-1 $\beta$  antiserum. This protein is not synthesized by the yeast transformed by the control vector pXKN18 (without IL-1 $\beta$  cassette). The secreted protein corresponds to the glycosylated form of IL-1 $\beta$ , which is demonstrated by the reduction of the apparent MW after treatment with the enzyme endo-N-acetylglucosaminidase H (Figure 11, lane 3). This host/vector pair *K. waltii*/pXPHO5, which is not yet optimized, secretes approximately 5 mg of IL-1 $\beta$  per liter of culture. The level of secretion of IL-1 $\beta$  by *K. waltii* is higher in the LPi medium than in the HPi medium, which suggests that the activity of the PHO5 promoter is regulated by the phosphate in *K. waltii*, as in *S. cerevisiae*.

## 6.2 Human serum albumin:

### 6.2.1 Construction of plasmid pYG140 (Figures 12-14).

A plasmid was constructed which comprises:

- an *E. coli* replicon,

- the gene aph under the control of the promoter k1 of the killer toxin of *K. lactis*

(EP361911) in which the HindIII site was eliminated by directed mutagenesis, and

- the gene bla, which confers resistance to ampicillin.

The gene aph, under the control of the promoter k1, is isolated from plasmid pKan707 (EP361991) in the form of a fragment PstI, which is cloned in the equivalent site of the phage M13mp7. The resulting plasmid is called pYG64 (Figure 12). The HindIII site present in this gene was destroyed by direct mutagenesis according to the method described by Taylor et al. (Nucl. Acid. Res. 13 (1985) 8749). The resulting plasmid is called pYG65. The oligodeoxynucleotide having served for the mutagenesis has the following sequence: 5'-GAAATGCATAAGCTCTTGCCATTCTCACCG-3', and it allows the transformation of the triplet CTT coding for leucine 185 in CTC. To construct plasmid pYG70, the part containing the bacterial replicon of vector pKan707 was isolated by digestion with the enzyme EcoRI and recircularization with T4 DNA ligase to obtain pYG69. The fragment PstI present in the latter vector containing gene aph was then replaced by the mutated equivalent fragment originating from pYG65. The resulting plasmid is called pYG70 (Figure 13).

This plasmid is then digested by EcoRI and re-ligated in the presence of an EcoRI-NarI-EcoRI adapter containing the following sequence: 5'-AATTGGCGCCG-3'.

The plasmid obtained is called pYG140 (Figure 14).

### 6.2.2 Introduction of an expression cassette for albumin (Figure 14)

The gene coding for preproSAH under the control of the promoter and the terminator of the gene PGK of *S. cerevisiae* was isolated in the form of an SalI-SacI fragment from an expression vector pYG19 (EP361991). This fragment was introduced into the corresponding sites of plasmid pYG140 to generate plasmid pYG141.

### 6.2.3 Construction of expression vector pYG142 (Figure 15)

Plasmids pYG141 and pKW1 are digested by the enzymes NarI and ClaI, respectively. After the ligation, 4 recombinant plasmids are obtained because of the existence of 2 forms, A and B, of pKW1, and of the orientation of the pKW1 part with respect to the pYG141 part.

Figure 15 describes the restriction map of one of these 4 plasmids: pYG142, containing the form B of pKW1.

The other plasmids are called pYG143, pYG144 and pYG145.

One sample of the strain *K. lactis* CBS 6430 was deposited at the CBS in Baarn (Netherlands) pursuant to the conditions of the Treatise of Budapest, on June 4, 1991 under the number CBS 290.91.

Table 1

Vectors	Fragment of pKW1	Shuttle vector	Marker
pBNA1	BglII-NheI 2.4Kb	Yip5 *BamHI-NheI	<u>URA3</u>
pNEA2	NheI-EcoRI 2.0kb	Yip5 NheI-EcoRI	<u>URA3</u>
pXXY2	XhoI-XbaI 0.55 kb	Yip5 SalI-NheI	<u>URA3</u>
pXXK3	XhoI-XbaI 0.55 kb	pKan21 SalI-XbaI	<u>Kan<sup>R</sup></u>
pBNB1/A3	BglIII-NheI 1.9 kb	Yip5 BamHI-NheI	<u>URA3</u>
pNEB1	NheI-EcoRI 2.5 kb	Yip5 NheI-EcoRI	<u>URA3</u>

Table 2

Vectors	Linearization site of pKW1	Shuttle vector	Marker
pKWS14 Pkwc11 PKWS 1	SalI ClaI SalI	pKan21*(SalI) pKan21 (AccI Yip5 (SalI)	Kan <sup>R</sup> Kan <sup>R</sup> URA3

Table 3. Transformation of *Saccharomyces cerevisiae* by vectors derived from pKW1

Vectors	Transformants Ura+ per µg of DNA
pKWS1	4 400
pBNA1	1 200
pBNB1/A3	7 600
pXXY2	4 000
pSK1	4 500

Table 4. Transformation of *Kluyveromyces waltii* by vectors derived from pKW1

Vectors	Replication support	Transformants G418-resistant per µg of DNA	
		pKW1 <sup>+</sup>	pKW1 <sup>-</sup>
pKWC11	pKW1 total	36 00 (98 %)	8 000 (100 %)
pXXK3	XbaI-XhoI	35 000 (49 %)	10 000 (2,8 %)
pKWS14	540bp de pKW1 pKW1 total	10 000 (92 %)	8 000 (29 %)

Table 5. Transformation of yeasts of the genus *Kluyveromyces* with the vector pKWC11

Species	Strain	GC%	Frequency of transformation per µg of DNA	Stability of the transformants (%)
<i>K. waltii</i>	CBS 6430	45.6	2400	100
<i>K. thermotolerans</i>	CBS 6340	46.2	4000	25

Claims

1. Plasmid pKW1 isolated from the strain *K. waltii* CBS6430, or any fragment or derivative thereof.
2. Fragment according to Claim 1, characterized in that it is a genetic element.
3. Plasmid characterized in that it comprises all or part of the sequence presented in Figure 3 or a derivative thereof.
4. Cloning and/or expression vector, characterized in that it comprises all or part of plasmid pKW1 of *K. waltii* CBS 6430 represented in Figure 1 or a derivative thereof.
5. Vector according to Claim 4, characterized in that it comprises a genetic element of at least plasmid pKW1.
6. Vector according to Claim 5, characterized in that it comprises the replication origin of plasmid pKW1.
7. Vector according to Claim 4, characterized in that it comprises the totality of plasmid pKW1.
8. Vector according to Claim 4, characterized in that it comprises all or part of the sequence presented in Figure 3 or a derivative thereof.
9. Vector according to Claim 4, 7 or 8, characterized in that plasmid pKW1 is linearized at the level of a functionally neutral restriction site.
10. Vector according to Claim 9, characterized in that it contains restriction site present on plasmid pKW1, or one which was artificially introduced to it.
11. Vector according to Claim 10, characterized in that it contains a restriction site artificially introduced into an intergenic region, and preferably in the region located between the genes B and D, or in the region located between gene D and IR2.
12. Vector according to Claim 10, characterized in that plasmid pKW1 is linearized at the level of the sites Clal(1), PstI(4608) or EcoRV(3072), the positions being given with reference to Figure 3.
13. Vector according to any one of Claims 4-12, characterized in that it contains, in addition, a heterologous DNA sequence comprising a structural gene which is at least under the control of signals allowing its expression.
14. Vector according to Claim 13, characterized in that the expression signals consist of one or more elements chosen from promoters, terminators, and secretion signals.
15. Vector according to Claim 14, characterized in that the promoters can be regulated.
16. Vector according to Claim 14, characterized in that the promoters are derived from yeast genes, and preferably from glycolytic yeast genes.
17. Vector according to Claim 13, characterized in that the structural gene codes for a polypeptide of pharmaceutical or agro-food interest.

18. Vector according to Claim 13, characterized in that the structural gene codes for a hybrid protein.

19. Vector according to Claim 13, characterized in that the structural gene(s) is (are) genes involved, at the genetic or biochemical level, in the biosynthesis of a metabolite.

20. Vector according to Claim 4, characterized in that it contains in addition an *E. coli* replicon.

21. Vector according to Claim 4, characterized in that it contains in addition at least one selection marker.

22. Recombinant cell containing a vector according to any one of Claims 4-21.

23. Cell according to Claim 22, characterized in that it is a yeast cell.

24. Cell according to Claim 23, characterized in that it is a yeast of the genus *Kluyveromyces* or *Saccharomyces*.

25. Method for the preparation of a polypeptide, characterized in that a recombinant cell is cultured according to one of Claims 22-24 and the produced polypeptide is recovered.

26. Method according to Claim 25, characterized in that the polypeptide is chosen from enzymes (such as, notably, superoxide dismutase, catalase, amylases, lipases, amidases, chymosine, etc.), blood derivatives (such as serum albumin, alpha- or beta-globin, factor VIII, factor IX, van Willebrand factor, fibronectin, alpha-1 antitrypsin, etc.), insulin and its variants, lymphokines (such as interleukins, interferons, colony stimulation factors (G-CSF, GM-CSF, M-CSF...), TNF, TRF, etc.), growth factors (such as growth hormone, erythropoietin, FGF, EGF, PDGF, TGF, etc.), apolipoproteins, or antigenic polypeptides for the preparation of vaccines (hepatitis, cytomegalovirus, Epstein-Barr, herpes, etc.).

27. Method according to Claim 26, characterized in that the polypeptide is human albumin or one of its variants or precursors.

28. Use of a recombinant cell according to one of Claims 22-24 as a catalyst in a bioconversion reaction.

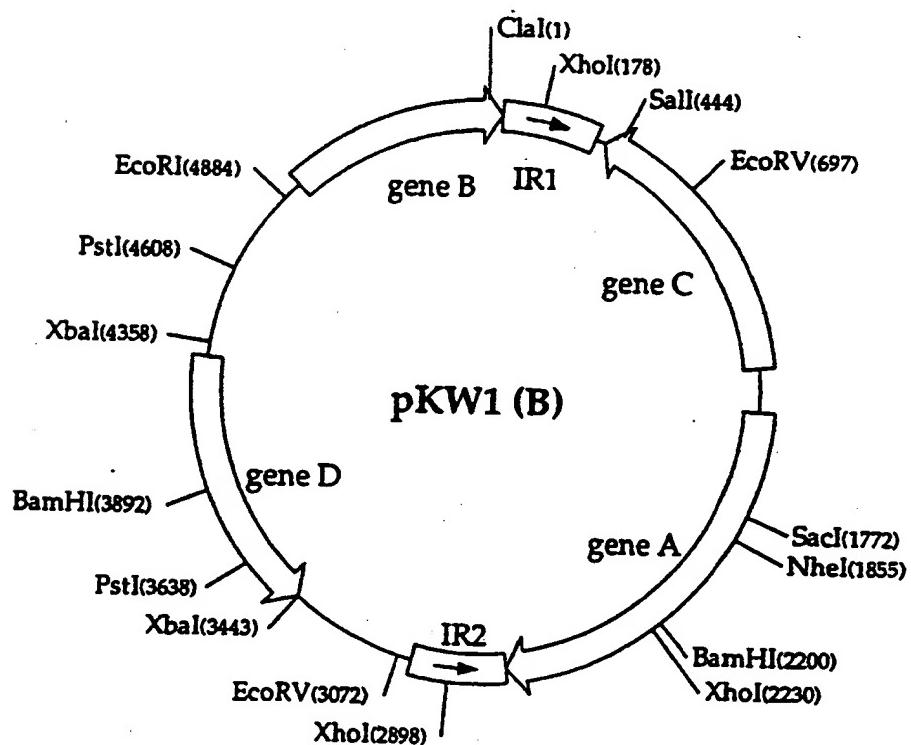
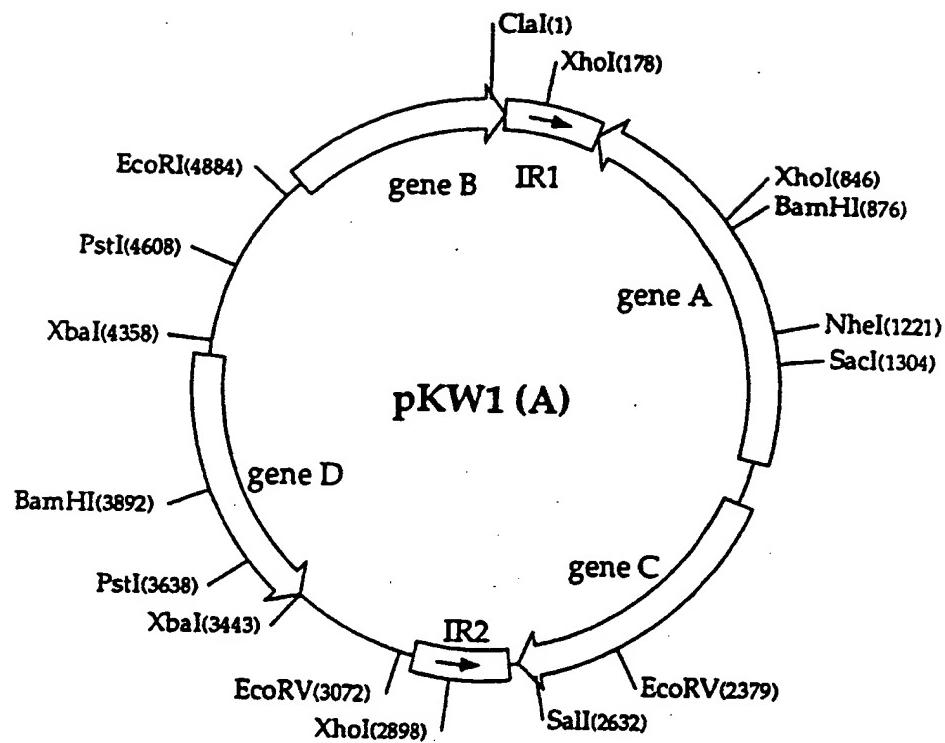


FIGURE 1

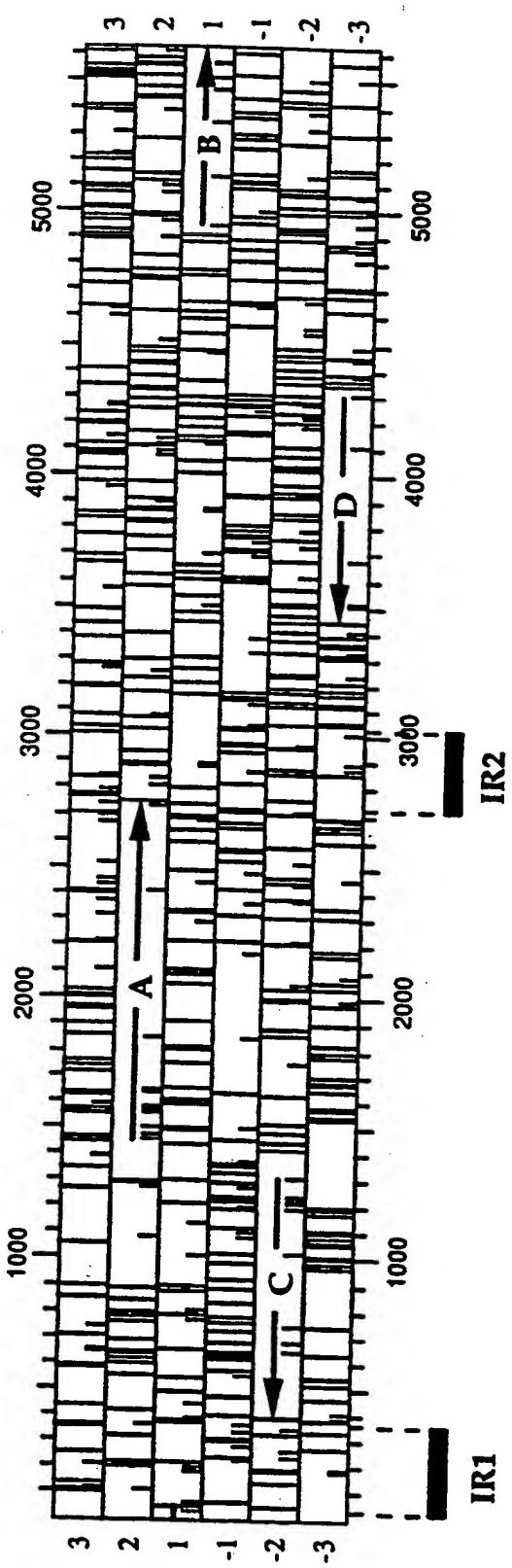


Figure 2

	1	10	20	30	40	50	60
1	ATCGATTGCC	AGCAACAGGG	TTTGCGCGT	ATGTCTTGT	TGGATACTAT	GTAAACAAAA	60
61	CAATCAATGT	ACAAAGAGCA	CAGCGGCCGG	CAGGTGGGAG	GACCCTCTGA	TGAGCCGGGG	120
121	ATATGGCGCT	CCCGCGCGTC	TAATATCCGG	ATTGGACTGG	AGGAGGACCA	AGGTTTCCTC	180
181	GAGGTCAATGT	GCCGCATCAT	GAGGGACGAA	AAATGGTAAG	GAATAGACCA	TTCCCTTACCA	240
241	TTTTTCGTCG	CCCGTGAGAT	TTTCCCATT	CCCCTTCCTT	ACCATTTTC	GTTCCGCACG	300
301	ACCCATCAC	ATTGTATATT	GAAATCTACT	TCCCTTCATA	CCGCTTGGCC	ATCGCGTAGG	360
361	AATGAACATA	TGTAGCCCT	CTAGTCCCT	ATGTTGTTCT	CAATCGACCC	TTCCATCGTG	420
421	AAGGGATCTC	CTCTTGAGGC	AATGTGACG	TTCGTACCCCT	CTTCCTACGG	GTCACGGGTT	480
481	CGTCTGTAAC	ACTTCTTCTT	GGTCTCCCTC	GGCCTCTTTT	ACTTGGCTGT	GGCCTGGTAT	540
541	TTTGAGTTC	ATCGAATAGG	CGCGGGTTT	CTTTTACTTG	AGGCCAGTCT	TGAATCTCTG	600
601	AACCGAGAGA	GGTTAGAGAA	GTTATTCTTT	TTAGATCATT	TAACATGTCC	TTGATGTCTT	660
661	CTGGCGTTGT	TTCGAACGCAC	ATCAGTCCCC	TTGAGCGATA	TCTGTCGTAA	TTGCCAGTA	720
721	TTACGTTGTG	CGCTTCTTCC	ATGTTTCAG	GGGCCACATG	CTTTAATGCC	TGCACTAAAA	780
781	CACTTCTGAC	AAGTGAACAA	TCTGCGAATG	CAGGCTCGTT	CGCCAATATC	AATCTCTGTA	840
841	GAGGAACACC	ACGATAGCTT	GTTGTGCCCT	TGCAGATAGT	GTTCAAGAAT	AGACTGAGGT	900
901	AGCCACGTAC	GTCGAGCTTT	ACAGACCTGG	TGCCAACGCC	AGAAAGAAGG	GCGAAGGGGT	960
961	CCTCTTCAAT	CGGAAAACGG	GTCAGGCTAT	CCACACGGTA	CTCTCTATAT	GTACTACGAG	1020
1021	TCATCGTCGT	GTACTTTGCG	AAGTGCAGGC	CGTGGCGATT	ATAGTCATC	TGGAAAGAAT	1080
1081	CCTGCTCATG	TGGCTCTTTC	TCCCCATTAT	CCTCTTCAGA	GTCAGCGAG	CCAGATGGCA	1140
1141	GCTCGGCTGG	CTCTCGCAGG	TACGTGGAT	CGGGCTCACC	TCGATCCACC	GTCATCTCAT	1200
1201	CAAGATTGCA	CATAATCTGC	ATCAGTCTCA	ACGTCAATAC	GACCGCGGAA	AAGGCGTCCG	1260
1261	AGTAGGAAAC	ATGCAGTTG	GAAAGGTCTA	ACAGCTTGAG	GCAGTGCATC	TTGGCCTCTC	1320
1321	TCTCTGTCAG	TCTTCTCTG	TTTGACCAA	TTTTCTCAA	GTGTATTACA	TTTGTCTTCTT	1380
1381	GGATTGCAAA	AATGTTGCGA	GTACTGCGTG	TTCTCGTTCA	AAAACAGCGT	TCGCTGAAAA	1440
1441	TTTAGGAGTT	CAGATGCAAC	GCGTAGTGCA	AATGGAGGAT	TCAAGTTGCA	GTAACAATAA	1500
1501	CATGGAACAC	CAAGGATCAG	TGTCGAGGA	GCTTATCTCC	AAAAATCTTA	TGAGCCTGAT	1560
1561	GGAAGAATTG	ATGTCTATGC	TCACTAATGA	GAAGGAGTTC	CAACGTGAAA	GGTCGCGTC	1620
1621	TCTCCTAGCC	TACATGATAA	TGGCTACTGG	TGAATTGGAA	GAGAAAAAGC	TCAGTACATT	1680
1681	TACCAAGTAT	TCCCGACGGA	TCAGGCAGAC	GGTAGAGTTC	GACAGCAACA	ACCAAATCGT	1740
1741	AAGATTGAG	TACCATTGAA	AGAATCCAC	AGAGCTCAAG	GAGACGCTGG	ACAAGGCCTT	1800
1801	AAACACTGTC	GTGTTTGAAA	TCAAGTCAA	AAAGAAGGTT	GTCTCCATGC	TGGAGCTAGC	1860
1861	TGCGAAGCTC	GACAAAAGGG	GATCAGATT	AGCGGGTGGT	ACGGTAGCTA	GTGAGGTCTC	1920
1921	GAAGCTGTG	CGGGAGGAAG	AAATTTGGCT	TCTTCTCGTG	AACGTGAAGA	ATACTATCCA	1980
1981	GGAAAAGGTG	CGCAAATCAT	CGCTAAGAGC	GGAAATTGACG	TATATTTGA	CAGCCTCATT	2040
2041	CTTCAATTGT	TGCAGACATA	GCGATCTCAG	GAACGCAGAC	CCCGCAACAT	TTGAGCTGGT	2100
2101	GCCAAATAAG	TATGTGGGCC	ACGTTGTCCTG	GGTTTGGTG	TGCGAGACCA	AGACCCGAAA	2160
2161	GCCGCGGGTTC	ATATACTTTT	TCCCTGTCAA	TACGGCCGCG	GATCCTCTAG	TAGCGCTTCA	2220
2221	TGATTTGTTG	TCGAGCACGT	TTCCCTTCAA	AAAGAGTCGG	ACGTCCGAAA	GAAAGCAGGA	2280

Figure 3 (a)

2281 ACAGGAATGG CAGATCGTTC GCGACGCATC 'AATCAACAAC TATGACCGGT TTGTTGGTAA 2340  
 2341 GCACGCTACG GAATCTGTCT TTGCCATCTT GCATGGTCCC AAATCACACT TGGGCCGGCA 2400  
 2401 CTTGATGAGT TCCTACTTGG CGTATAACCCA CCATGGGAA TGGGTCTCAC CATATGGGAA 2460  
 2461 TTGGTCAGCT GGGAAAGGAA CCATTGAAAG CAGCGTGGCA AGGGCCAAGT ACGCACATGT 2520  
 2521 TCAAGCCGAG ATCCCAAGCG ATCTTTCGC CTTCTGTCT CAGTACTATC AGGAATCAA 2580  
 2581 ATCGGGCGAT TTCGAGCTTA ACGACACCAAG CAAAGACCCA ACAAAAGCTGG TACGGCACTC 2640  
 2641 GGCTAGTCAT CTGGAAATCA ATCGAACCTA TGGTCCATGG AGTAGATTGG TTAACAAGGA 2700  
 2701 TGTTTTAGGC TTTGTTCATC CCTACGCGAT GGCCAAGCGG TATGAAGGGAA AGTAGATTTC 2760  
 2761 AATATACAAT GTGATAGGGT CGTGCAGAAC GAAAATGGT AAGGAACGGG AAATGGGAAA 2820  
 2821 ATCTCACGGG CGACGAAAAA TGGTAAGGAA TGGTCTATTTC CTTACCATTTC TTGTCCTCCTC 2880  
 2881 ATGATGCGGC ACATGACCTC GAGGAAACCT TGGTCCTCCT CCAGTCCAAT CCGGATATTA 2940  
 2941 GACGCGCGGG AGCAGCCTAT CCCCGGCTCA TCAGAGGGTC CTCCCACCTG CCGGCCGCTG 3000  
 3001 TGCTCTTGT ACATTGATTG TTTTGTTTA GTATTACCTG ACAATCATT TATATTTGT 3060  
 3061 CAGTATTTCT TGATATCTGG CCCAAATGAA AGTACAAATA CAAGTACAAG TACAACACCT 3120  
 3121 ATTCAATTGTT ACCGTATCTC TATACTATTAA TCCCTTATTTC CTGCCTAATT AACTACTTTC 3180  
 3181 TACCGGTGCG TTCTTAAAGC TGAGTAGGCA CTTCAGCCCCG AAGTGAATAT CGCTGGCCTC 3240  
 3241 ATTATCTGTT TGGTGTATGG CCAATGAAAC TCCAACCGAT TTGATCCACT TGCAAGGACC 3300  
 3301 GTAGGATAGA CTTATTTTA CCATCCCATT ATCTGTGCAG TGGACACCTC TGTTTATCTT 3360  
 3361 TACCCGCCCT TCAGAGAACAA TAATCTGTCT TACAAAAGAG TAGACCTCAT TTTGGCGAG 3420  
 3421 AAGCCTCTTC TGTTTGACCT CATCTAGAAC CTTTTGGGCT CTCCCAGCGC AATCAGATAG 3480  
 3481 TCTTGAGA GTGTCCTCTT CGTCACCATG CTCATCAGAA GGAGCTGGTT GCGTTCCAAT 3540  
 3541 TGAGAAAGCT CGTCCCAAAA GGCTGCATGG TCTATAGGTG GTCCCGAACCG CGAGGATTCA 3600  
 3601 CCTTGAGACT CATTGTCCAC AGTAGGCTCC TGATCCTCTG CAGCAGCCGC TTGCTGAATG 3660  
 3661 GATTCTGCTC TTGACATTAC CATTTCGAGA AGCCATAGTC TTGGATGGT TCTAGCAGAT 3720  
 3721 CTTTCAACAT CCAAACGTGA ACCGAACCAT TGCCGACAAC TTATTCTAAC CTCAAGCTTT 3780  
 3781 GAGAGGTTCT CATACTTGGT TTGTGCAGCC TCAAGATCCG TAAGTCGGT GAAGAAATCT 3840  
 3841 AAAGCTTCGC GTTTGGACC TAACCGAATG ATGATTGGGT GCCGTTTCG TGGATCCAGA 3900  
 3901 AATGGTGTAA GCTCCTCTAT ACTATCATTC GTGGGACACT CTATCCCTTC CCCAATATGC 3960  
 3961 TTCTTTACAA TAGAAGGTAG CTGCGTATAC TTGTTCCGAA CAAACAAAGAT GTGACTCTCT 4020  
 4021 AGCCGCTTTA CTATTGTGTT CAACACAACG TAGGGTTTG GCCACTCAA TGTTAATTCT 4080  
 4081 GATCGAGGTG CTGATGCTGT GTTCTTGTG ATCATGTATG AATAGTAGAA CAATGCCAAT 4140  
 4141 TGAGGATGCT CACTGTATTTC CTTAGGCAGT TTGGTGCTGC CATAAGGTGA CACCAGCTCC 4200  
 4201 TTTAGCTGTG GAACATAGAG TACATCAACA GTTCTACTA GACAACATTC CTTGAGACAA 4260  
 4261 TTATGATAGC CATCATGGTC GGTTATCTTA GGTATTTTT GAGCTTCACT CATCTTCGAA 4320  
 4321 GCAACCGTGA TAGATTCGAT TGAAGTTCAA ATTCTTATCT AGATGGGTGA TTTGTTTATC 4380  
 4381 ATAATTTCACA ATACAGTCTG TTTTAATTG CTCGAAGTTG CAGTGAAAGA TGAAAAAAGG 4440  
 4441 GGCTTCATTT TGCACTACAG ACTTACCTG ATGTAAGGGAA TTTTCATCAT AAAAGCAATT 4500  
 4501 TCTCGTATAC AAAATGCAGT TTGTTTCGTA TACAGCAAAG CATAGAATAC ACTGTACACT 4560  
 4561 GCATTTTAC TACAAGAAAA GTTTTTTTG CTGCTGTGAC TGGAAATGCTG CAGCAGTACC 4620

Figure 3 (b)

4621 TAGATAGAAA AATGGCATAA AACTGAAATT TTATAGTCAT TTTTCGTGTC TTTCAATTCAA 4680  
4681 TTTTTCTCG CAAAAGTTT CTACAAAAGC AGTCAAAATT GCAATAAGTA TACACTAATT 4740  
4741 ACAAGGCAGTC TGAGCGCGTG ACTTGAGCGC GTGACGTAAT CGCGAGCTAC GAAAGTTGTT 4800  
4801 TGGGCCTCG ACATCGGATC GACAGAAGAG GTAAGAATAT TGGGAAAGTA CATTCAATTAA 4860  
4861 CCACAACAAT CGAGAGATTA GTGGAATTCA GTCATTAATG AAAGGTAGGG TAGTCCGCCT 4920  
4921 ACTCTTAGTT CTACATTCAAG AATACGCATG CAGTCCTCAA GTTCTGATGA AGACGACCTA 4980  
4981 ATTGACCCCTA TAATTCATCC CAAATCGTTC TATAGGGCGG CTAACGAGAT ACCGAGAGAC 5040  
5041 TTTTGGTTG CGATCCCCAT CAGCGCCTAT GTTTTAGCG TATTTGCTAA ATCAGTACGA 5100  
5101 GATGACTTAC AGGGGCATT AACGGCGCGA GATATGGCAT TAGCTTATCG TGAACGGCAG 5160  
5161 TACTTCACA GACGCTGGGA GACACGAAAC GACCAGCTTG AGATCCCAGA CTGGTCTGAG 5220  
5221 ATCCCAGAAC GGTCTCTCGG GTTACTGGAT CGCCCTCCTT GTATCACTGT GGATCTAGCA 5280  
5281 AGAGAACTGC GCGAACTATC TCAAAAATGG ATCGGAGCAT TCGATCTGGG ATCGAAGATG 5340  
5341 TCTGGCAGGC TTCTTCTACA GCTTCTGTAC ACCCAGCTGT CATGCCAAA TGAGGCTGTT 5400  
5401 TTCAATAAGC TTTACTGCCT CGTCAAACTA CTAAACAAGG ACGTAAATCG TGGGGACCGA 5460  
5461 GCCCTTATGG ACTCTGTATT GAGACCCTT TTTGTCGAGA ATCCATACAT GGGTGAACTA 5520  
5521 GATGAAGAAA TACTTGATAA GATATGGTCC AATTGACTG AAATGAGAAG TCAAGAGTGG 5580  
5581 AAACGTATAG CGGAAGCGTT GTCAGGCGAG AATAATGAC

5619

Figure 3 (c)

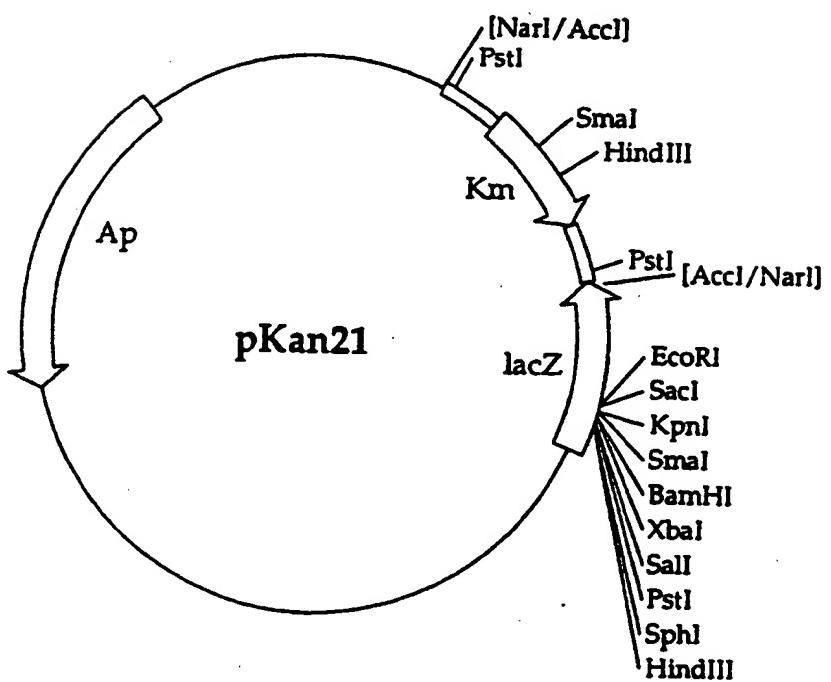
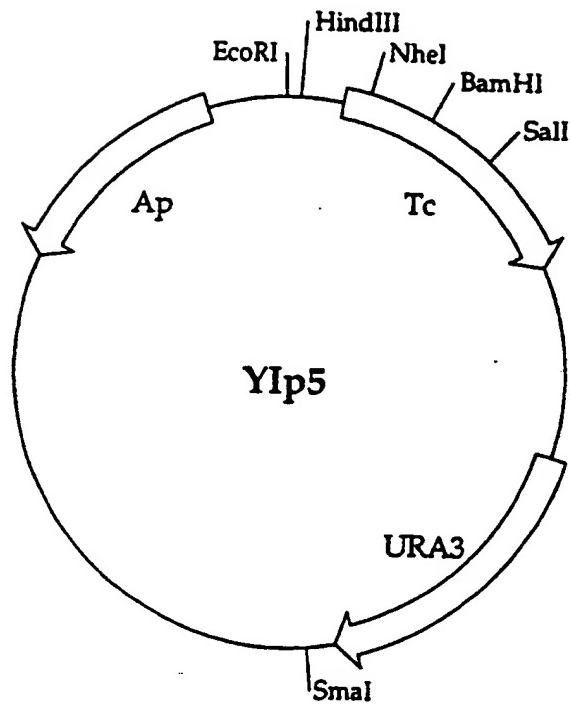


FIGURE 4

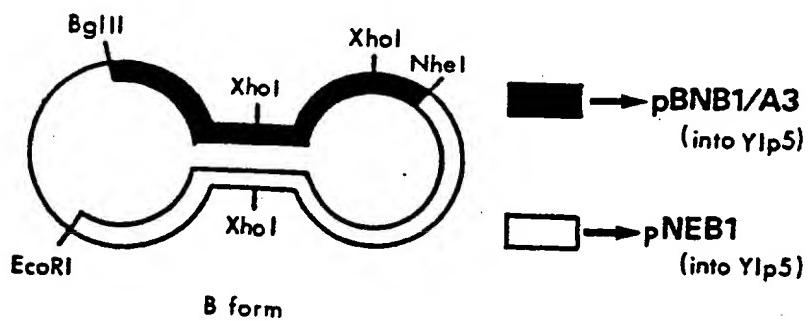
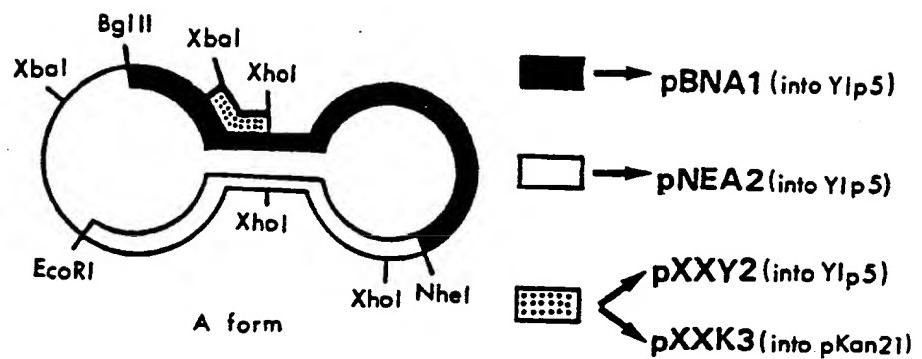


FIGURE 5

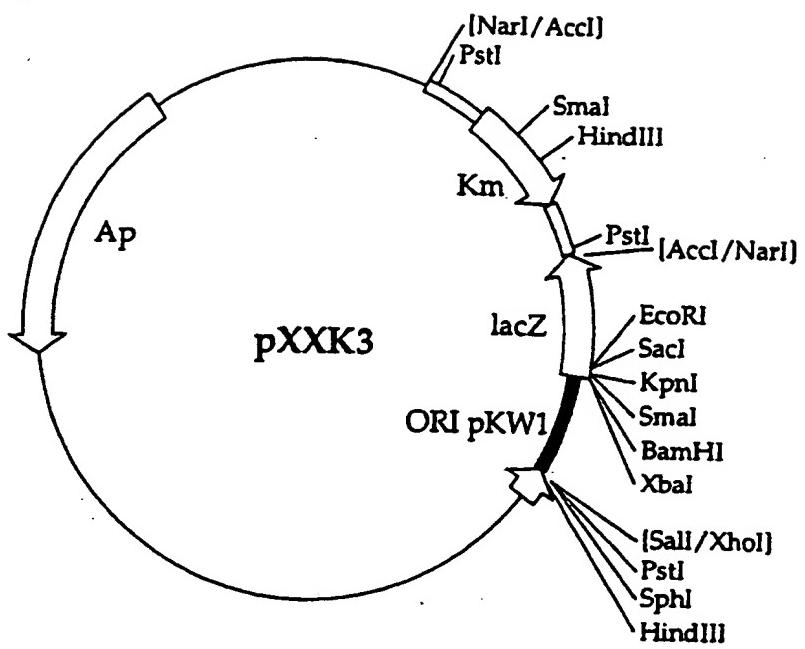


Figure 6

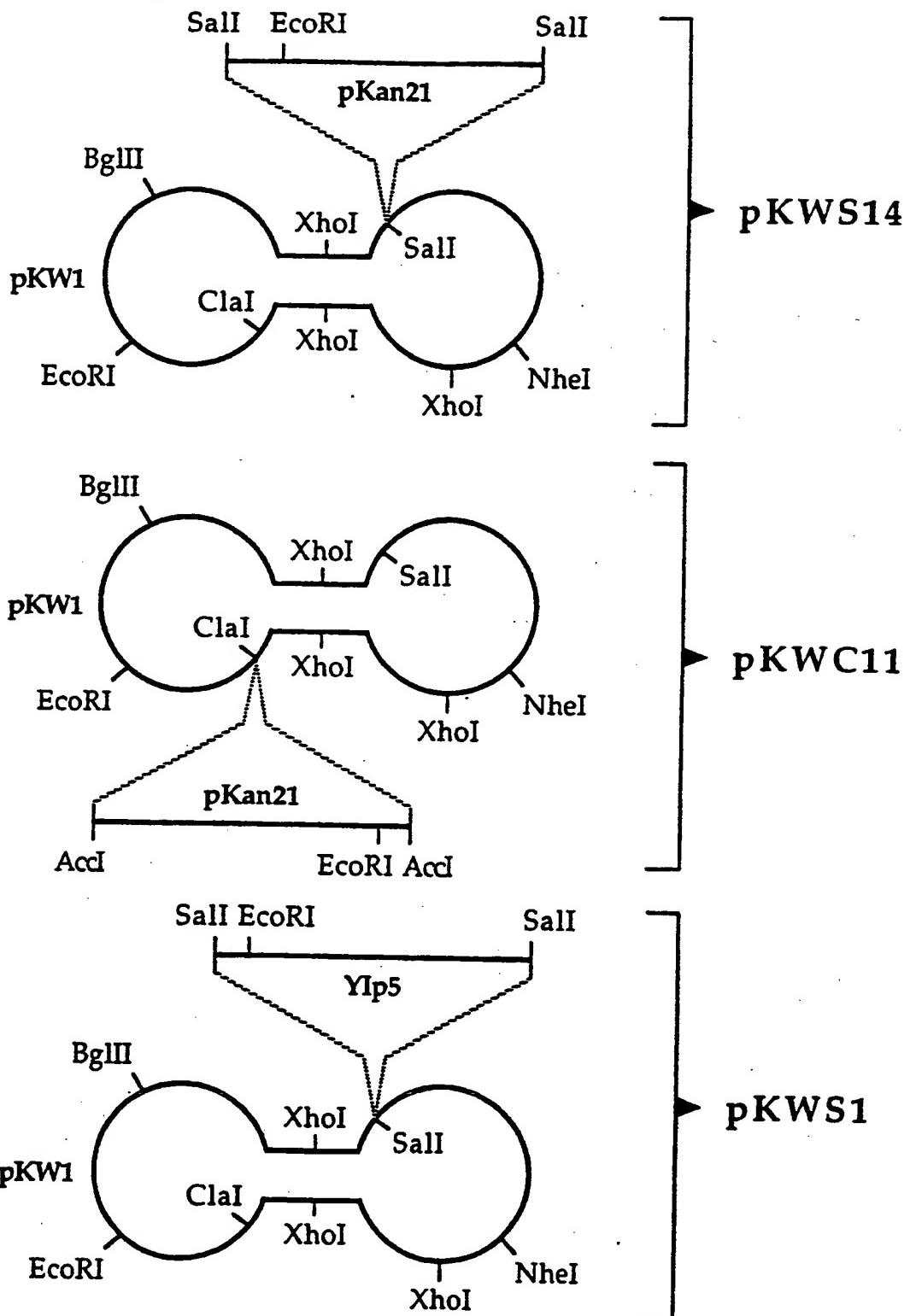


Figure 7

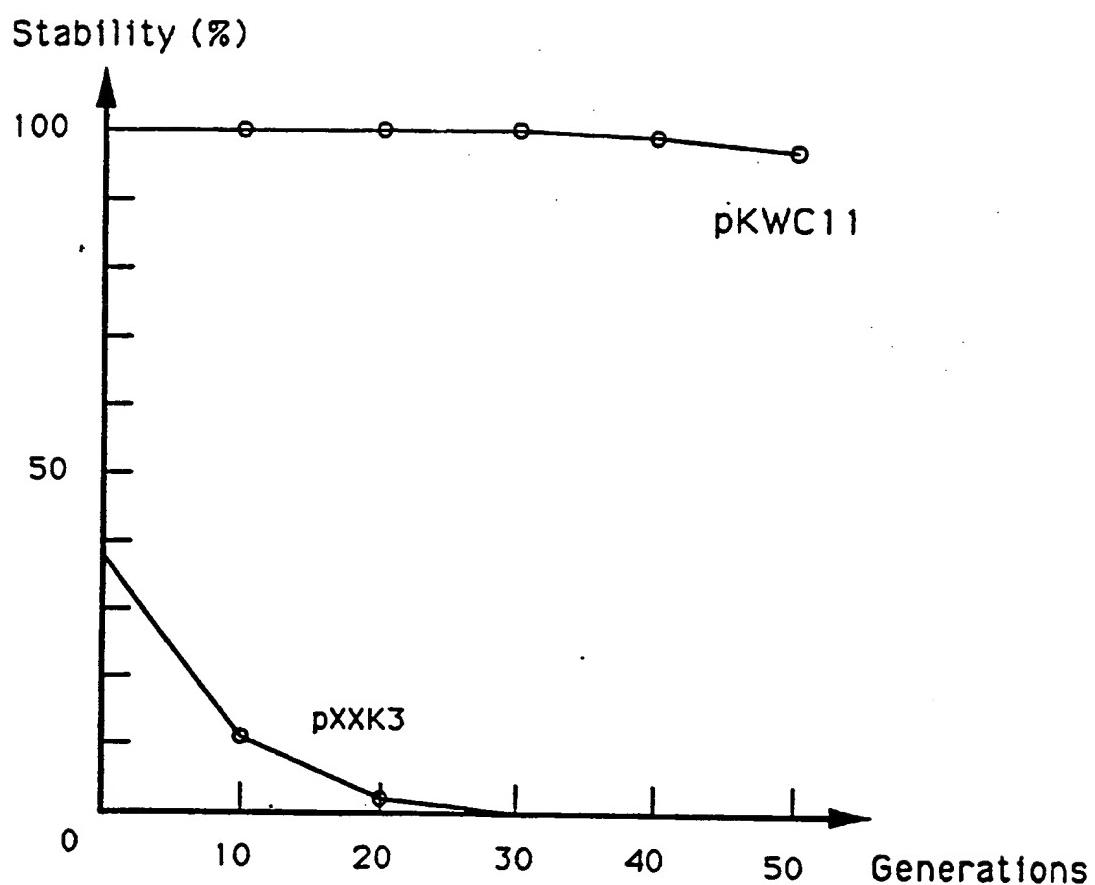


Figure 8

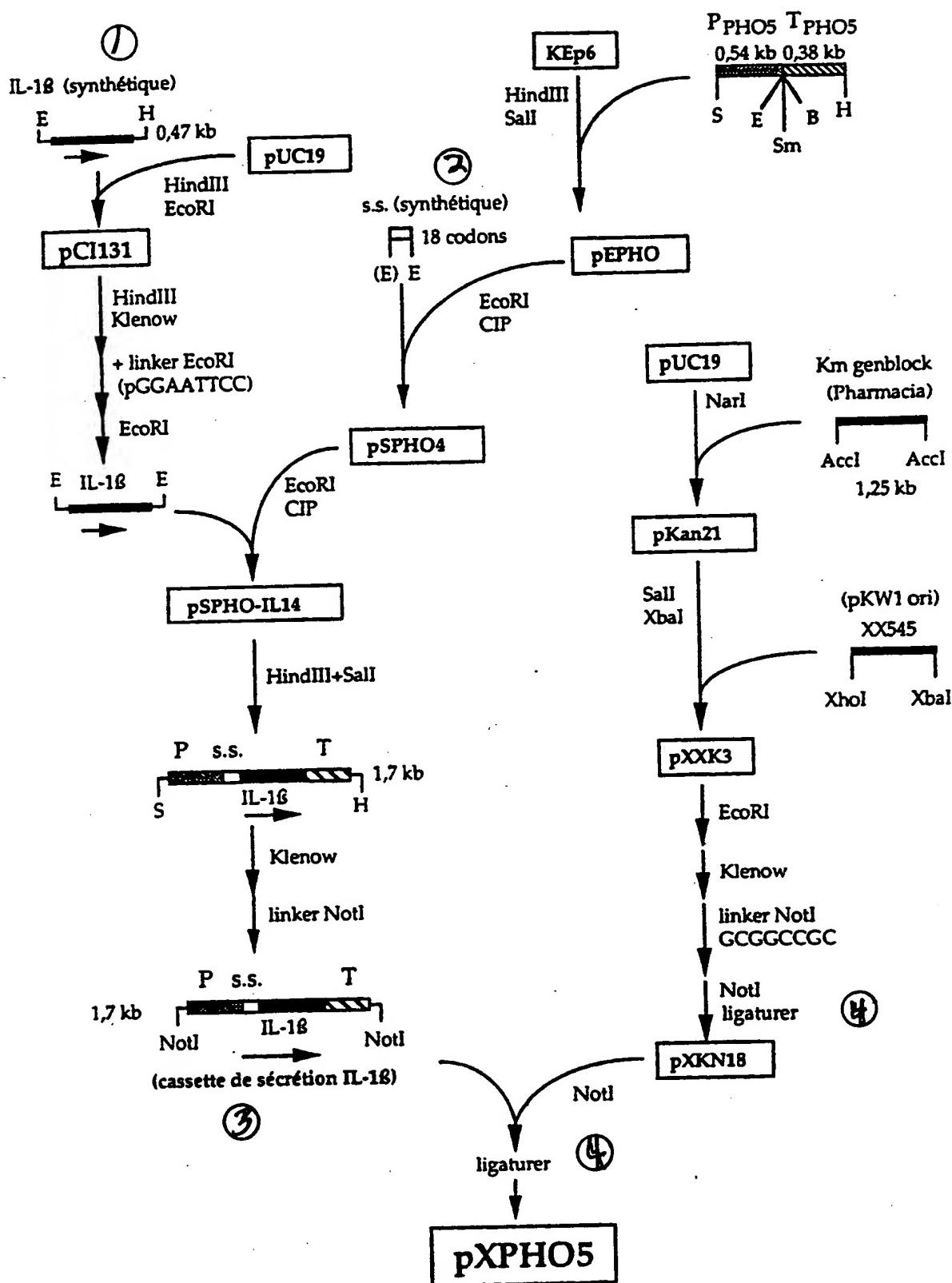


FIGURE 9

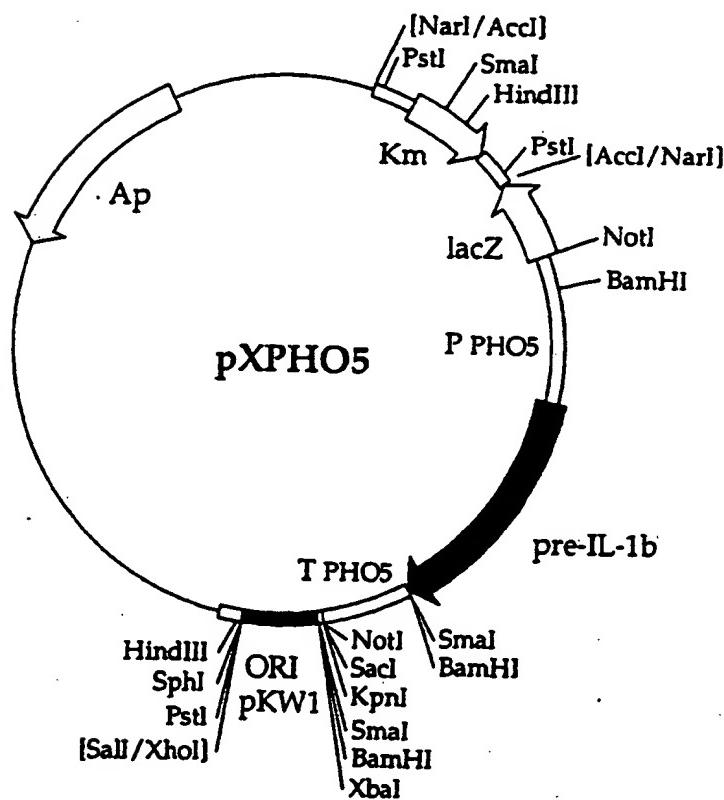
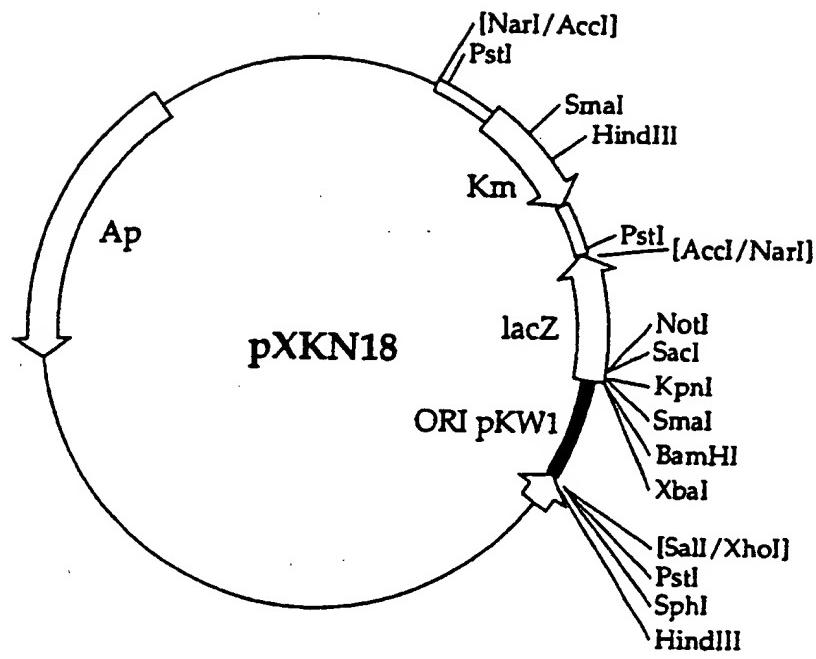
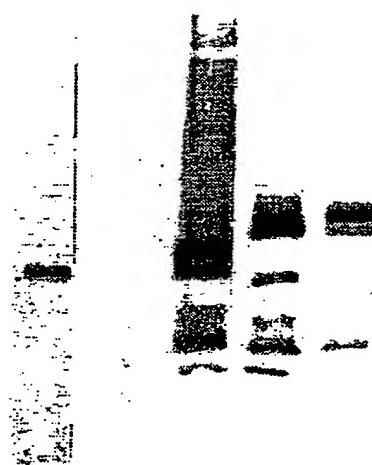


FIGURE 10

13 / 17

1 2 3 4 5

44.2 ►  
27.8 ►  
18.3 ►  
15.3 ►



FEUILLE DE REMplacement

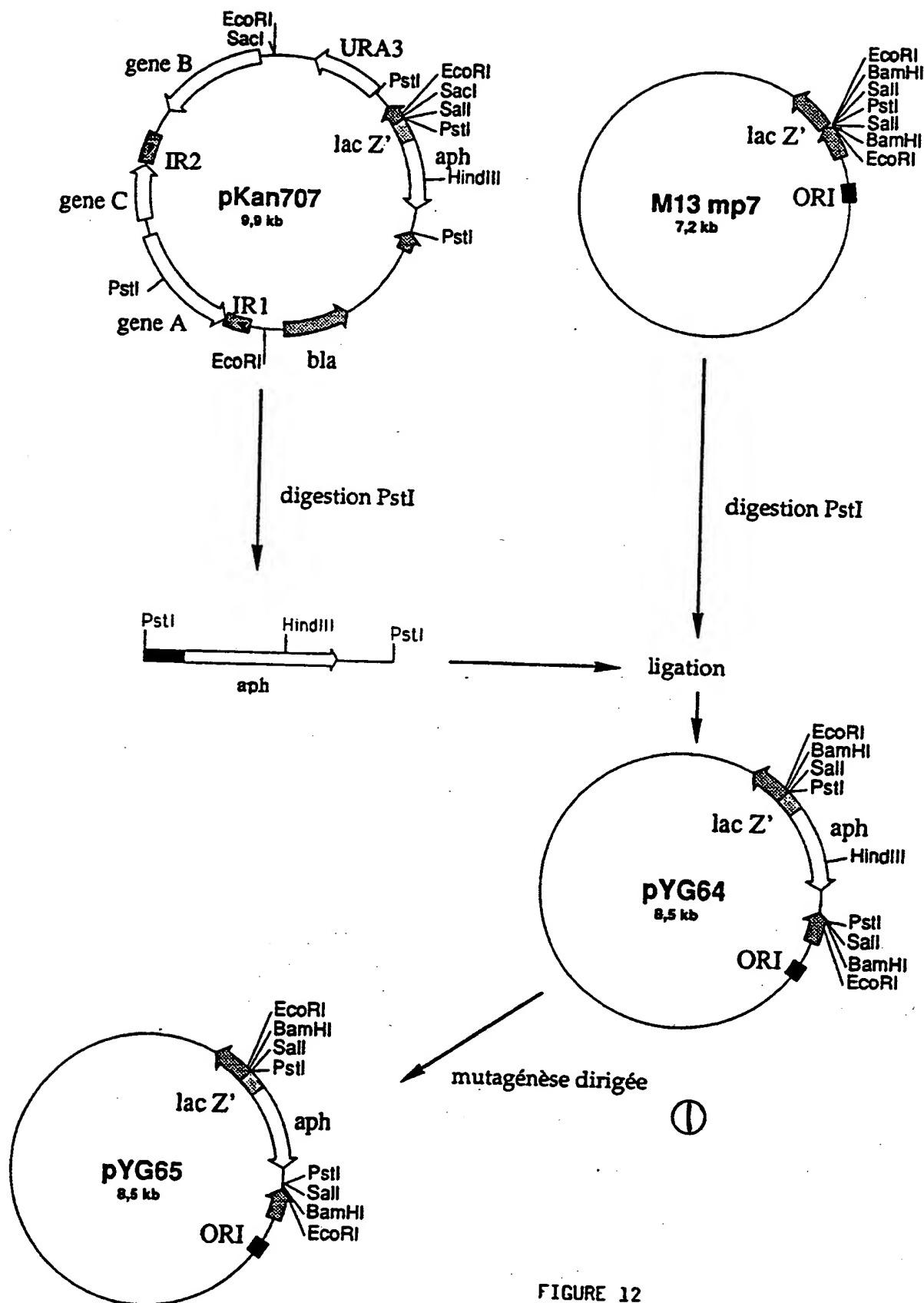


FIGURE 12

Key: 1 Directed mutagenesis

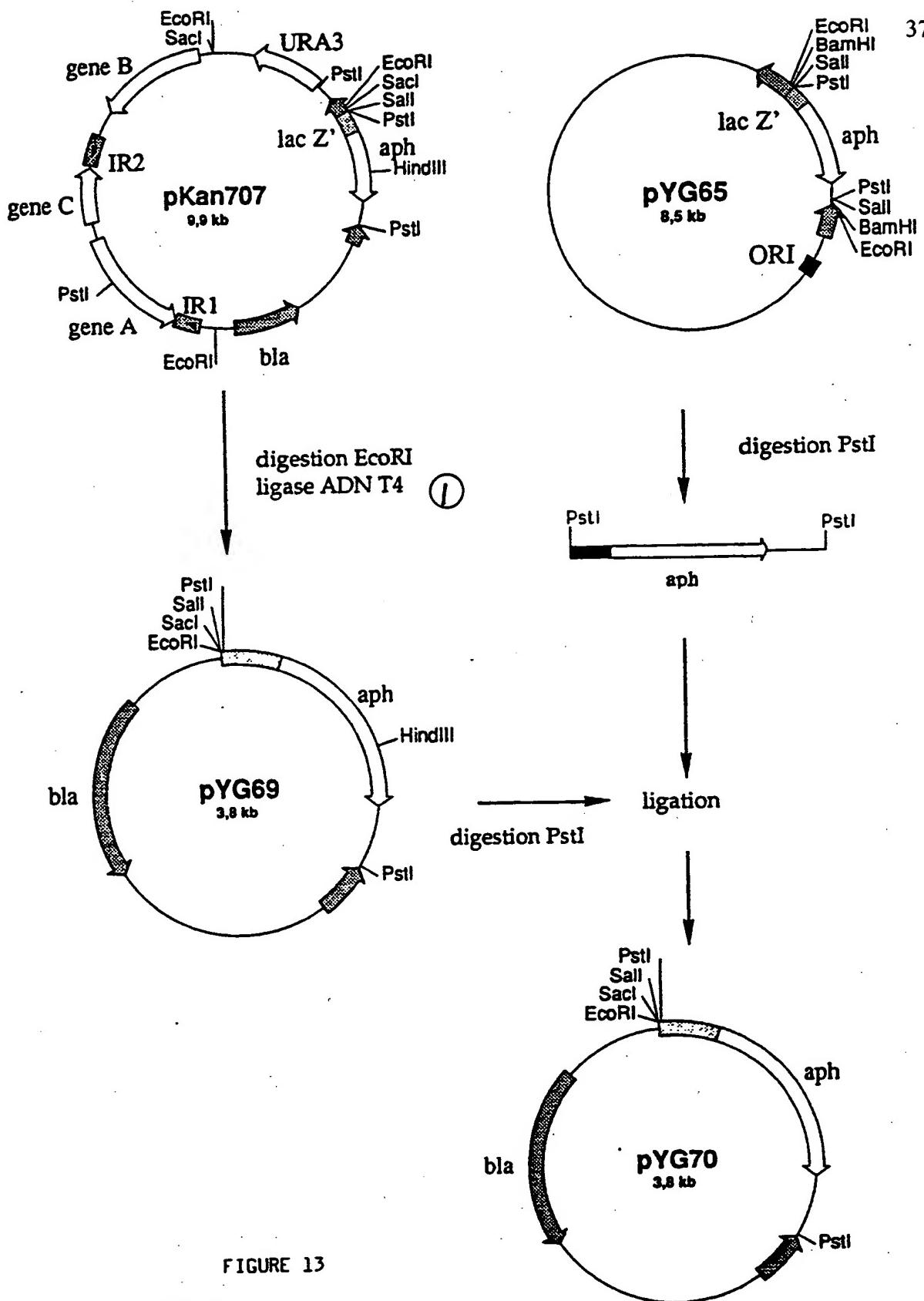


FIGURE 13

Key: 1 T4 DNA ligase

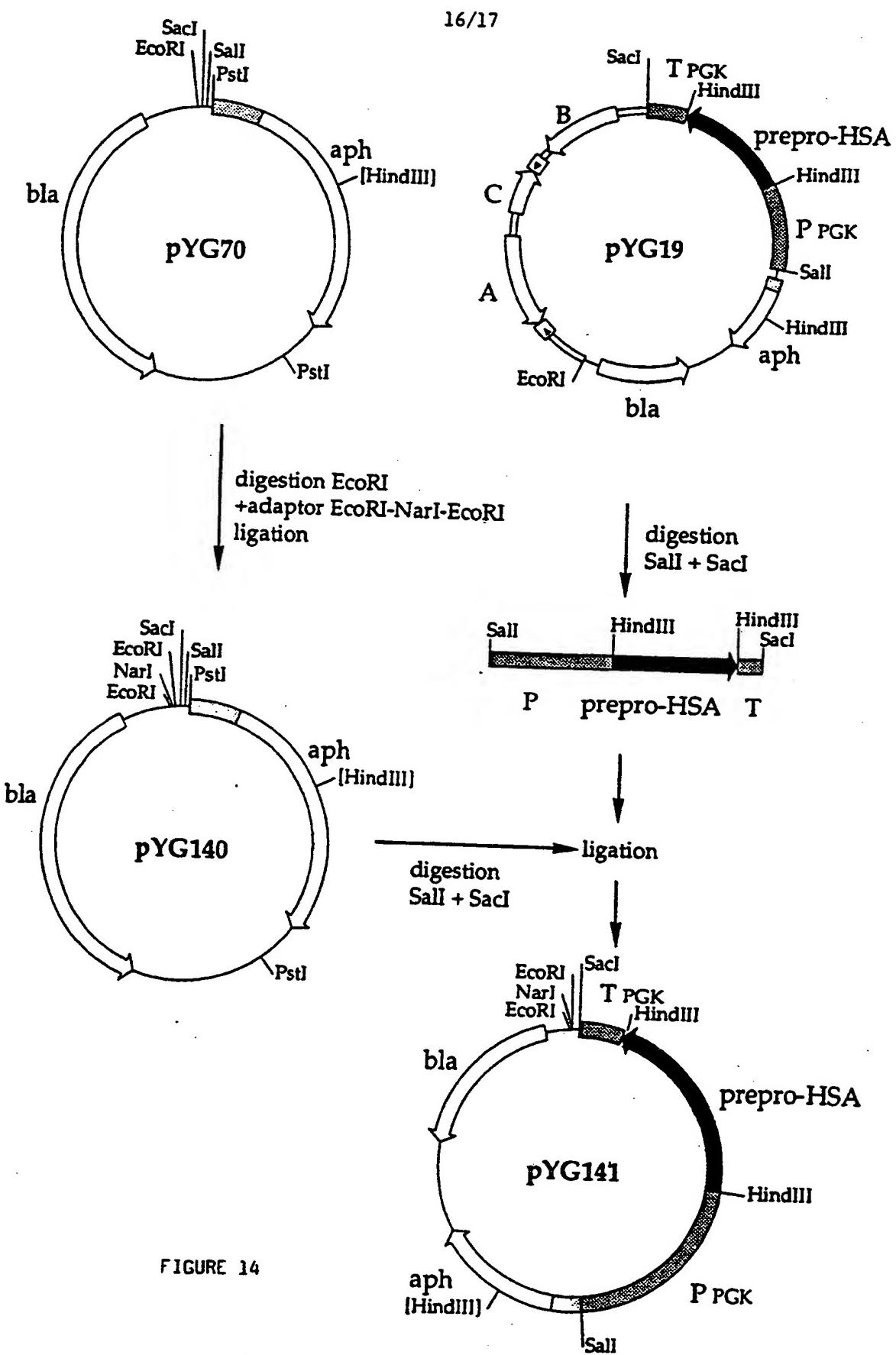


FIGURE 14

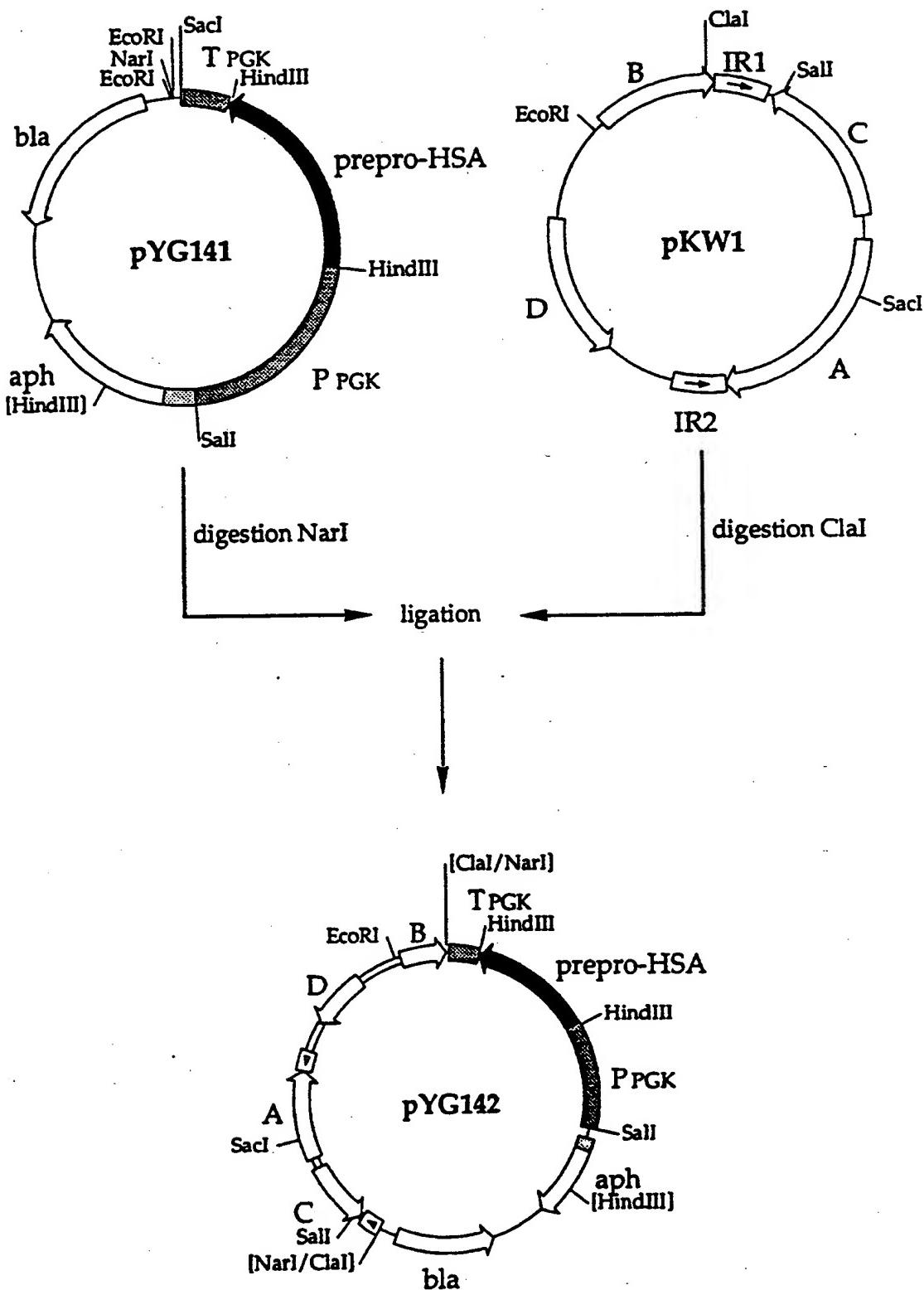


FIGURE 15

International Application No.: PCT/[illegible]

## MICROORGANISMS

Optional sheet pertaining to the microorganism mentioned on page 20, lines 10-11 of the description<sup>1</sup>:

**A. Identification of the deposition<sup>2</sup>:**

Other depositions are identified on a supplementary sheet<sup>3</sup>.

**Name of the deposition institute<sup>4</sup>:**

Centraalbureau voor Schimmelcultures (CBS)

**Address of the deposition institute (including the zip code and the country<sup>4</sup>):**

Oosterstraat 1, P. O. Box 273 – 3740 AG Baarn (Netherlands)

**Date of deposition<sup>5</sup>:**

June 4, 1991

**Sequential No.<sup>6</sup>:**

CBS 290.91

**B. Supplemental indications<sup>7</sup> (to be filled out only if necessary). A separate sheet is attached for the continuation of this information**

With regard to the designations under which a European patent is filed, a sample of the microorganism deposited will not be accessible prior to the date of publication of the mention of the grant of the European patent or the date when the application is rejected, withdrawn or considered to be rejected, unless a sample is made available by an expert designated by the requester (rule 28(4) of the EPC).

**C. Designated states for which the indications are given<sup>3</sup> (if the indications are not given for all the designated states)**

**D. Information supplied separately<sup>8</sup> (to be filled out if necessary)**

The information listed below will be later subjected to International Office<sup>9</sup> (specify the general nature of the indications, for example, "Sequential No. of the deposit"

E. The present sheet was received with the international application when it was filed (to be verified by the receiving office)

[signature]

N. Saada

(Authorized employee)

Date received (originating from the applicant) by the International Office<sup>11</sup>

(Authorized employee/officer)

Form PCT/RO/134 (January)

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/FR 92/00559

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>		
Int.Cl.5: C12N15/81		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols)		
Int.Cl.5: C12N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	FR,A,2 635 115 (RHONE-POULENC SANTE) 9 February 1990 —	
X,P	JOURNAL OF GENERAL MICROBIOLOGY vol. 138, No.2, February 1992, COLCHESTER pages 337 - 345 X.J. CHEN ET AL. 'Characterization circular plasmid from the yeast Kluyveromyces waltii' —	1-28
A	EP,A,0 301 670 (GIST-BROCADES N.V.) 1 February 1989 — —/—	
<input type="checkbox"/> Further documents are listed in the continuation of Box C.		<input type="checkbox"/> See patent family annex.
<ul style="list-style-type: none"> <li>* Special categories of cited documents:</li> <li>"A" document defining the general state of the art which is not considered to be of particular relevance</li> <li>"B" earlier document but published on or after the international filing date</li> <li>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</li> <li>"O" document referring to an oral disclosure, use, exhibition or other means</li> <li>"P" document published prior to the international filing date but later than the priority date claimed</li> </ul>		
Date of the actual completion of the international search 9 October 1992 (09.10.92)		Date of mailing of the international search report 26 October 1992 (26.10.92)
Name and mailing address of the ISA/ European Patent Office Facsimile No.		Authorized officer  Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/FR 92/00559

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CURRENT GENETICS vol. 19, No. 3, 1991, BERLIN, GERMANY pages 163 - 167 C. WILSON AND H. FUKUHARA 'Distribution of mitochondrial r1-type introns and the associated open reading frame in the yeast genus Kluyveromyces'	

**ANNEX TO THE INTERNATIONAL SEARCH REPORT  
ON INTERNATIONAL PATENT APPLICATION NO. FR 9200559  
SA 61252**

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information. 09/10/92

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
FR-A-2635115	09-02-90	AU-B-	623425	14-05-92
		AU-A-	3933289	08-02-90
		EP-A-	0361991	04-04-90
		FR-A-	2649991	25-01-91
		JP-A-	2276589	13-11-90
-----	-----	-----	-----	-----
EP-A-0301670	01-02-89	US-A-	4943529	24-07-90
		AU-A-	2014888	02-03-89
		JP-A-	2000476	05-01-90
-----	-----	-----	-----	-----